

Recovery rates in investment-grade pools of credit assets: A large deviations analysis

Konstantinos Spiliopoulos^a, Richard B. Sowers^{b,*}

^a Division of Applied Mathematics, Brown University, Providence, RI 02912, United States

^b Department of Mathematics, University of Illinois at Urbana–Champaign, Urbana, IL 61801, United States

Received 1 January 2011; received in revised form 9 August 2011; accepted 11 August 2011

Available online 23 August 2011

Abstract

We consider the effect of recovery rates on a pool of credit assets. We allow the recovery rate to depend on the defaults in a general way. Using the theory of large deviations, we study the structure of losses in a pool consisting of a continuum of types. We derive the corresponding rate function and show that it has a natural interpretation as the favored way to rearrange recoveries and losses among the different types. Numerical examples are also provided.

© 2011 Elsevier B.V. All rights reserved.

MSC: 60F05; 60F10; 91G40

Keywords: Recovery rates; Default rates; Credit assets; Large deviations

1. Introduction

Understanding the behavior of *large pools of credit assets* is currently a problem of central importance. Banks often hold such large pools and their risk-reward characteristics need to be carefully managed. In many cases, the losses in the pool are (hopefully) *rare* as a consequence of diversification. In [16], we have used the theory of large deviations to gain some insight into several aspects of rare losses in pools of credit assets. Our interest here is the effect of *recovery*. While a creditor either defaults or doesn't (a Bernoulli random variable), the amount recovered may in fact take a continuum of values. Although many models assume that recovery rate is

* Corresponding author.

E-mail addresses: kspiliop@dam.brown.edu (K. Spiliopoulos), r-sowers@illinois.edu (R.B. Sowers).

constant—i.e., a fixed deterministic percentage of the par value, in reality the statistics of the amount recovered should be a bit more complicated. The statistics of the recovered amount should depend on the number of defaults; a large number of defaults corresponds to a bear market, in which case it is more difficult to liquidate the assets of the creditors. Our goal is to understand how to include this effect in the study of rare events in large pools. We would like to look at these rare events via some ideas from statistical mechanics, or more accurately the theory of *large deviations*. Large deviations formalizes the idea that nature prefers “minimum energy” configurations when rare events occur. We would like to see how these ideas can be used in studying the interplay between default rate and recovery rate.

Our work is motivated by the general challenge of understanding the effects of nonlinear interactions between various parts of complicated financial systems. One of the strengths of the theory of large deviations is exactly that it allows one to focus on propagation of rare events in networks. Our interest here is to see how this can be implemented in a model for recovery rates which depend on the default rate.

This work is part of a growing body of literature which applies the theory of large deviations to problems of rare losses in credit assets and tries to approximate the tail distribution of total losses in a portfolio consisting of many positions. For a survey on some existing large deviations methods applied to finance and credit risk, see [14]. In [8,13] rare event asymptotics for the loss distribution in the Gaussian copula model of portfolio credit risk and related importance sampling questions are studied. In [9] the author considers saddle point approximations to the tails of the loss distribution. Measures such as Value-at-Risk (VaR) and expected shortfall have been developed in order to characterize the risk of a portfolio as a whole, see for example [2,3,6,7,10,12]. Issues of recovery have also been considered in the works [1,15] and in the references therein. Our work is most closely related to [4,11], where the dynamics of a configuration of defaults was studied. In [4] it is assumed that a “macro-environmental” variable Y , to which all the obligors react, can be chosen so that conditional on Y , the recovery and default rates in a pool of finite number of “types” are independent. In [11] the recovery rate is assumed to be independent of the defaults in a pool of one type with a brief treatment of a pool with finite types. Our work here is explicitly interested in the *dependence* of the recovery rate on the fraction of the number of defaults and the framework of our efforts is a continuum of types. The case of a continuum of types requires slightly more topological sophistication. The dependence assumption and the continuum of types unavoidably complicate the proof since several auxiliary technical results need to be proven.

The novelty of our result stems from the following two things. The first novelty is that the distribution of the recoveries for the defaulted assets depends upon the number of defaults in a fairly general way. This allows for consideration of the case when recovery rates are affected by the number of defaults. Note however that the individual defaults are assumed independent. The second novelty is more technical in nature. In particular, we prove the large deviations principle for the joint family of random variables (L_N, ν_N) where L_N is the average loss and ν_N is the empirical measure on type-space determined by the names which default. In addition to the large deviations principle, we derive various equivalent representations for the rate function which give some more insight to the favored way to rearrange recoveries and losses among the different types and also ease its numerical computation.

The paper is organized as follows. In Section 2 we introduce our model and establish some notation. In Section 3 we study the “typical” behavior of the loss of our pool; we need to understand this before we can identify what behavior is “atypical”. In Section 4, we present some formal sample calculations and our main results, Theorems 4.6 and 4.8. These calculations are

indicative of the range of possibilities and they illustrate the main results. In Section 5 we present some numerical examples which illustrate some of the main aspects of our analysis and conclude with a discussion and remarks on future work. The proof of Theorem 4.6 is in Sections 6–8. Section 9 contains an alternative expression of the rate function, which is a variational formula which optimizes over all possible configuration of recoveries and defaults, and which leads to a Lagrange multiplier approach which can be numerically implemented. Lastly, Section 10 contains the proofs of several necessary technical results.

The model at the heart of our analysis is in fact very stylized. Since our primary interest is the interaction between default rates and recovery rates, our model focuses on this effect, but simplifies a number of other effects. In particular, we assume that the defaults themselves are independent. Our work complements the existing literature and hopefully contributes to our understanding of the interplay and interaction of recovery and default rates. It seems plausible that more realistic models (e.g., which include a systemic source of risk) can be analyzed by techniques which are extensions of those developed here.

2. The model

In this section we introduce our model and explain our goals. Let us start by considering a single bond (or “name”). For reference, let us assume that all bonds have par value of \$1. If the bond defaults, the assets underlying the bond are auctioned off and the bondholder recovers r dollars, where $r \in [0, 1]$. We will record the default/survival coordinate as an element of $\{0, 1\}$, where 1 corresponds to a default and 0 to survival.

The pool suffers a loss when a bond defaults, and the amount of the loss is $\$1 - r$, where $r \in [0, 1]$ is the recovery amount (in dollars). Define

$$\Delta_n \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if the } n\text{th name has defaulted} \\ 0 & \text{otherwise} \end{cases}$$

and let ℓ_n be the non-zero loss that occurs in the event of default of the n th name, i.e., when $\Delta_n = 1$. Equivalently, ℓ_n is the non-zero exposure of the n th name that occurs when $\Delta_n = 1$. Clearly if r_n is the recovery of the n th name then $\ell_n = 1 - r_n$.

The default and loss rates in the pool are then

$$D_N \stackrel{\text{def}}{=} \frac{1}{N} \sum_{n=1}^N \Delta_n \quad \text{and} \quad L_N \stackrel{\text{def}}{=} \frac{1}{N} \sum_{n=1}^N \ell_n \Delta_n. \quad (2.1)$$

We denote by $p^{N,n}$ the risk-neutral probability of default of the n th name in a pool of N names. Moreover, let $q^{N,n}(D_N, dr)$ be the distribution of the recovery of the n th name in a pool of N names. Notice that we allow $q^{N,n}(D_N, dr)$ to depend on the default rate D_N .

Let us make now the aforementioned discussion rigorous by introducing the probability space that we consider. The minimal state space for a single bond is the set $E_\circ \stackrel{\text{def}}{=} \{0\} \cup (\{1\} \times [0, 1])$. Since we want to consider a pool of bonds, the state space in our model will be $E \stackrel{\text{def}}{=} E_\circ^{\mathbb{N}}$, where $\mathbb{N} \stackrel{\text{def}}{=} \{1, 2, \dots\}$.

E_\circ , as a closed subset of $\{0, 1\} \times [0, 1]$ is also Polish, and thus E is also Polish. We endow E with the natural σ -algebra $\mathcal{F} \stackrel{\text{def}}{=} \mathcal{B}(E)$.

The only remaining thing to specify is a probability measure on (E, \mathcal{F}) . For each $N \in \mathbb{N}$, fix $\{p^{N,n} : n \in \{1, 2, \dots, N\}\} \subset [0, 1]$. These are the risk-neutral default probabilities

of the names when the pool has N names. We next fix $\{\varrho^{N,n} : n \in \{1, 2, \dots, N\}\} \subset C([0, 1]; \mathcal{P}[0, 1])^1$; i.e., a collection of probability measures on $[0, 1]$, the range of the recovery, indexed by the default rate D_N . Namely, for each $N \in \mathbb{N}$ and $n \in \{1, 2, \dots, N\}$, $\varrho^{N,n}$ is a probability measure that depends on the default rate D_N . Some concrete examples are in Section 5.1.

For each $n \in \mathbb{N}$ and $\omega = (\omega_1, \omega_2, \dots)$, define the coordinate random variable $X_n(\omega) = \omega_n$. For each $N \in \mathbb{N}$, we then fix our risk-neutral probability measure $\mathbb{P}_N \in \mathcal{P}(E)$ (with associated expectation operator \mathbb{E}_N) by requiring that

$$\mathbb{E}_N \left[\prod_{n=1}^N f_n(X_n) \right] = \mathbb{E}_N \left[\prod_{n=1}^N \left\{ (1 - p^{N,n}) f_n(0) + p^{N,n} \int_{r \in [0,1]} f_n(1, r) \varrho^{N,n}(D_N, dr) \right\} \right]$$

for all $\{f_n\}_{n=1}^N \subset B(E_o)$. In particular, the aforementioned construction implies that for each $n \in \mathbb{N}$, $\{\Delta_n\}_{n=1}^N$ is an independent collection of random variables with $\mathbb{P}_N\{\Delta_N = 1\} = p^{N,n}$ under \mathbb{P}_N .

With this probabilistic structure in place, we will clearly want to be able to condition on the default rate so that we can then focus on the recovery rates. To this end we define the σ -algebra

$$\mathcal{D} \stackrel{\text{def}}{=} \sigma\{\Delta_n : n \in \mathbb{N}\}.$$

We are interested in the *typical behavior* and *rare events* in this system. We seek to understand the structure of these rare events in our model. We are interested in the following questions:

- What is the typical behavior of the system? In other words we want to characterize $\bar{L} \stackrel{\text{def}}{=} \lim_{N \rightarrow \infty} L_N$.
- What can we say about rare events in this general setting? In particular we compute the asymptotics of $\mathbb{P}_N\{L_N \geq l\}$ as $N \rightarrow \infty$, particularly for $l > \bar{L}$. Then $\{L_N \geq l\}$ is an “atypical” event.

Understanding the structure of rare events may give guidance in how to optimally control how rare events propagate in large financial networks.

Remark 2.1. Our formulation allows for a fairly general dependence of the recovery distribution on the default rate. This is also partially exploited in Section 5.1 where we study several concrete examples. On the other hand, the general treatment of this paper unavoidably complicates the proof of the law of large numbers and especially of the large deviations principle.

3. Typical events: a law of large numbers

Let us start our analysis by identifying the “typical” behavior of L_N as $N \rightarrow \infty$ defined by Eq. (2.1). In order to motivate the calculations, we first investigate how D_N and L_N behave as $N \rightarrow \infty$ in a homogeneous pool, Example 3.1, and in a heterogeneous pool of two types, Example 3.2.

Example 3.1 (Homogeneous Example). Fix $p \in [0, 1]$ and $\varrho \in C([0, 1]; \mathcal{P}[0, 1])$ and let $p^{N,n} = p$ and $\varrho^{N,n} = \varrho$ for all $N \in \mathbb{N}$ and $n \in \{1, 2, \dots, N\}$.

¹ We shall write $\varrho \in C([0, 1]; \mathcal{P}[0, 1])$ as a map from $[0, 1] \times \mathcal{B}[0, 1]$ into $[0, 1]$ such that for each $D \in [0, 1]$, $A \mapsto \varrho(D, A)$ is an element of $\mathcal{P}[0, 1]$ and such that $D \mapsto \varrho(D, \cdot)$ is weakly continuous.

Let us see what D_N and L_N look like in this setting. By the law of large numbers $\lim_{N \rightarrow \infty} D_N = p$. Thus in distribution

$$L_N \approx \frac{1}{N} \sum_{1 \leq n \leq pN} (1 - \xi_n)$$

where the ξ_n 's are i.i.d. with distribution $q(p, \cdot)$. We should consequently have that

$$\lim_{N \rightarrow \infty} L_N = p \int_{r \in [0,1]} (1-r) q(p, dr). \quad \square$$

Example 3.2 (Heterogeneous Example). Fix p_A and p_B in $[0, 1]$ and fix q_A and q_B in $C([0, 1]; \mathcal{P}[0, 1])$. Every third bond will be of type A and have default probability p_A and recovery distribution governed by q_A , and the remaining bonds will have default probability p_B and recovery distribution governed by q_B . In other words, the rate of default among the A bonds is p_A and the rate of default among the B bonds is p_B .

For future reference, let us separate the defaults and into the two types. If we define

$$D_N^A \stackrel{\text{def}}{=} \frac{1}{\lfloor N/3 \rfloor} \sum_{\substack{1 \leq n \leq N \\ n \in 3\mathbb{N}}} \Delta_n \quad \text{and} \quad D_N^B \stackrel{\text{def}}{=} \frac{1}{N - \lfloor N/3 \rfloor} \sum_{\substack{1 \leq n \leq N \\ n \notin 3\mathbb{N}}} \Delta_n$$

we have

$$D_N = \frac{\lfloor N/3 \rfloor}{N} D_N^A + \frac{N - \lfloor N/3 \rfloor}{N} D_N^B \approx \frac{1}{3} D_N^A + \frac{2}{3} D_N^B.$$

Thus $\lim_{N \rightarrow \infty} D_N^A = p_A$ and $\lim_{N \rightarrow \infty} D_N^B = p_B$, so

$$\lim_{N \rightarrow \infty} D_N = \bar{D} \stackrel{\text{def}}{=} \frac{p_A}{3} + \frac{2p_B}{3}.$$

Thus we should roughly have

$$L_N \approx \frac{1}{N} \left\{ \sum_{1 \leq n \leq p_A N/3} (1 - \xi_n^A) + \sum_{1 \leq n \leq 2p_B N/3} (1 - \xi_n^B) \right\}$$

where the ξ_n^A 's have law $q_A(\bar{D}, \cdot)$, the ξ_n^B 's have distribution $q_B(\bar{D}, \cdot)$ and the ξ_n^A 's and ξ_n^B 's are all independent. Consequently, it seems natural that

$$\lim_{N \rightarrow \infty} L_N = \frac{p_A}{3} \int_{r \in [0,1]} (1-r) q_A(\bar{D}, dr) + \frac{2p_B}{3} \int_{r \in [0,1]} (1-r) q_B(\bar{D}, dr),$$

the first term being the limit of the losses from the type A names, and the second term being the limit of the losses from the type B names. \square

In view of our examples, it seems reasonable that we should be able to describe the average loss in the pool in terms of a frequency count of

$$\gamma^{N,n} \stackrel{\text{def}}{=} (p^{N,n}, q^{N,n}).$$

We note that $\gamma^{N,n}$ takes values in the set $\mathbf{X} \stackrel{\text{def}}{=} [0, 1] \times C([0, 1]; \mathcal{P}[0, 1])$. Since $\mathcal{P}[0, 1]$ is Polish, so is $C([0, 1]; \mathcal{P}[0, 1])$, and thus \mathbf{X} is Polish. We will henceforth refer to elements of \mathbf{X}

as types. For each $N \in \mathbb{N}$, we now define $U_N \in \mathcal{P}(X)$ as

$$U_N \stackrel{\text{def}}{=} \frac{1}{N} \sum_{n=1}^N \delta_{\gamma^{N,n}}.$$

Assumption 3.3 is our main standing assumption. It is a natural assumption since it makes the family $\{U_N\}_{N \in \mathbb{N}}$ relatively compact. It will be assumed throughout the paper, even though this may not be always stated explicitly.

Assumption 3.3. We assume that $U \stackrel{\text{def}}{=} \lim_{N \rightarrow \infty} U_N$ exists (in $\mathcal{P}(X)$). \square

Remark 3.4. In the case of [Example 3.1](#), we have that $U = \delta_{(p,q)}$, while in the case of [Example 3.2](#), we have that $U = \frac{1}{3}\delta_{(p_A, q_A)} + \frac{2}{3}\delta_{(p_B, q_B)}$.

We can now identify the limiting behavior of L_N . Define

$$\begin{aligned} \bar{D} &\stackrel{\text{def}}{=} \int_{\gamma=(p,q) \in X} p U(d\gamma) \quad \text{and} \\ \bar{L} &\stackrel{\text{def}}{=} \int_{\gamma=(p,q) \in X} \left\{ p \int_{r \in [0,1]} (1-r) q(\bar{D}, dr) \right\} U(d\gamma). \end{aligned} \tag{3.1}$$

To see that these quantities are well-defined, note that

$$(p, q) \mapsto p \quad \text{and} \quad v \mapsto \int_{r \in [0,1]} (1-r) v(dr) \tag{3.2}$$

are continuous mappings from, respectively, X and $\mathcal{P}[0, 1]$, to $[0, 1] \subset \mathbb{R}$. The continuity of the first map of (3.2) implies that \bar{D} is well-defined. Combining the continuity of both maps of (3.2), we get that the map

$$(p, q) \mapsto p \int_{r \in [0,1]} (1-r) q(\bar{D}, dr)$$

is a continuous map from X to $[0, 1] \subset \mathbb{R}$; thus \bar{L} is also well-defined.

Lemma 3.5. Let [Assumption 3.3](#) hold and consider \bar{L} given by (3.1). Then, for each $\varepsilon > 0$, we have that

$$\lim_{N \rightarrow \infty} \mathbb{P}_N \{ |L_N - \bar{L}| \geq \varepsilon \} = 0. \quad \square$$

Proof. For $q \in C([0, 1]; \mathcal{P}[0, 1])$ and $D \in [0, 1]$, let us first define

$$\Gamma(D, q) \stackrel{\text{def}}{=} \int_{r \in [0,1]} (1-r) q(D, dr).$$

Again we can use (3.2) and show, by the same arguments used to show that \bar{D} and \bar{L} of (3.1) are well-defined, that Γ is well-defined, and furthermore that it is continuous on

$[0, 1] \times C([0, 1]; \mathcal{P}[0, 1])$. For each $N \in \mathbb{N}$, define

$$\hat{L}_N \stackrel{\text{def}}{=} \frac{1}{N} \sum_{n=1}^N \Delta_n \Gamma(D_N, q^{N,n}), \quad \bar{D}_N \stackrel{\text{def}}{=} \frac{1}{N} \sum_{n=1}^N p^{N,n} = \int_{\gamma=(p,q) \in X} p \mathbf{U}_N(d\gamma)$$

$$\bar{L}_N \stackrel{\text{def}}{=} \frac{1}{N} \sum_{n=1}^N p^{N,n} \Gamma(\bar{D}, q^{N,n}) = \int_{\gamma=(p,q) \in X} p \Gamma(\bar{D}, q) \mathbf{U}_N(d\gamma);$$

note that \hat{L}_N is a random variable but \bar{D}_N and \bar{L}_N are deterministic. Note also that by weak convergence,

$$\lim_{N \rightarrow \infty} \bar{L}_N = \bar{L} \quad \text{and} \quad \lim_{N \rightarrow \infty} \bar{D}_N = \bar{D}.$$

The claim will follow if we can prove that

$$\lim_{N \rightarrow \infty} \mathbb{E}_N \left[|L_N - \hat{L}_N|^2 \right] = 0 \quad \text{and} \quad \lim_{N \rightarrow \infty} \mathbb{E}_N \left[|\hat{L}_N - \bar{L}_N|^2 \right] = 0. \quad (3.3)$$

To see the first part of (3.3), we calculate that

$$L_N - \hat{L}_N = \frac{1}{N} \sum_{n=1}^N (\ell_n - \Gamma(D_N, q^{N,n})) \Delta_n.$$

Conditioning on $\mathcal{D} = \sigma\{\Delta_n : n \in \mathbb{N}\}$, we have that

$$\mathbb{E}_N \left[\left(L_N - \hat{L}_N \right)^2 \middle| \mathcal{D} \right]$$

$$= \frac{1}{N^2} \sum_{n=1}^N \left\{ \int_{r \in [0,1]} \left((1-r) - \Gamma(D_N, q^{N,n}) \right)^2 q^{N,n}(D_N, dr) \right\} \Delta_n^2 \leq \frac{1}{N}.$$

This implies the first part of (3.3).

To see the second part of (3.3), we write that $\hat{L}_N - \bar{L}_N = \mathcal{E}_N^1 + \mathcal{E}_N^2$ where

$$\mathcal{E}_N^1 = \frac{1}{N} \sum_{n=1}^N [\Gamma(D_N, q^{N,n}) - \Gamma(\bar{D}_N, q^{N,n})] \Delta_n$$

$$\mathcal{E}_N^2 = \frac{1}{N} \sum_{n=1}^N \Gamma(\bar{D}_N, q^{N,n}) [\Delta_n - p^{N,n}].$$

We first calculate that

$$\mathbb{E}_N \left[\left| \mathcal{E}_N^2 \right|^2 \right] = \frac{1}{N^2} \sum_{n=1}^N \left(\Gamma(\bar{D}_N, q^{N,n}) \right)^2 p^{N,n} (1 - p^{N,n}) \leq \frac{1}{4N}.$$

Thus $\lim_{N \rightarrow \infty} \mathbb{E}_N \left[|\mathcal{E}_N^2|^2 \right] = 0$. Similarly,

$$\mathbb{E}_N \left[|D_N - \bar{D}_N|^2 \right] = \frac{1}{N^2} \sum_{n=1}^N p^{N,n} (1 - p^{N,n}) \leq \frac{1}{4N}.$$

Therefore

$$\lim_{N \rightarrow \infty} \mathbb{E}_N \left[|D_N - \bar{D}|^2 \right] = 0.$$

To bound \mathcal{E}_N^1 , fix $\eta > 0$. Due to [Assumption 3.3](#), Prohorov's theorem implies that $\{\mathbf{U}_N; N \in \mathbb{N}\}$ is tight, so there is a $K_\eta \subset\subset C([0, 1]; \mathcal{P}[0, 1])$ such that

$$\sup_{n \in \mathbb{N}} \mathbf{U}_N([0, 1] \times K_\eta^c) < \eta.$$

Defining

$$\omega_\eta(\delta) \stackrel{\text{def}}{=} \sup_{\substack{\varrho \in K_\eta \\ D_1, D_2 \in [0, 1] \\ |D_1 - D_2| < \delta}} |\Gamma(D_1, \varrho) - \Gamma(D_2, \varrho)|$$

for all $\delta > 0$, compactness of K_η and $[0, 1]$ and continuity of Γ imply that $\lim_{\delta \searrow 0} \omega_\eta(\delta) = 0$. Thus

$$\begin{aligned} |\mathcal{E}_N^1| &\leq \mathbb{E}_N \left[\frac{1}{N} \sum_{n=1}^N \left| \Gamma(D_N, \varrho^{N,n}) - \Gamma(\bar{D}_N, \varrho^{N,n}) \right| \right] \\ &= \mathbb{E}_N \left[\int_{\gamma=(p, \varrho) \in X} |\Gamma(D_N, \varrho) - \Gamma(\bar{D}_N, \varrho)| \mathbf{U}_N(d\gamma) \right] \\ &\leq 2\eta + 2\omega_\eta(\delta) + \tilde{\mathbb{P}}_N \{ |D_N - \bar{D}| \geq \delta \}. \end{aligned}$$

Take $N \nearrow \infty$, then let $\delta \searrow 0$. Finally let $\eta \searrow 0$. We conclude that indeed $\lim_{N \rightarrow \infty} \mathbb{E}_N[|\mathcal{E}_N^1|] = 0$. Hence, the second part of (3.3) also holds which concludes the proof of the lemma. \square

4. Problem formulation and main results

Let us now set up our framework for considering atypical behavior of the L_N 's; i.e., large deviations. We motivate the main result, [Theorem 4.6](#) through a formal discussion related to the setting of [Examples 3.1](#) and [3.2](#). The proof of [Theorem 4.6](#) is in Sections 6–8. In Section 6 we prove compactness of the level sets of the rate function. In Sections 7 and 8 we prove the lower and upper limits of large deviations respectively. An equivalent insightful representation of the rate function is given in [Theorem 4.8](#) and its proof is in Section 9. This representation is a variational formula which optimizes over all possible configuration of recoveries and defaults, and which leads to a Lagrange multiplier approach which can be numerically implemented.

For the convenience of the reader we here recall the concept of the large deviations principle and the associated rate function.

Definition 4.1. If X is a Polish space and \mathbb{P} is a probability measure on $(X, \mathcal{B}(X))$, we say that a collection $(\xi_n)_{n \in \mathbb{N}}$ of X -valued random variables has a large deviations principle with rate function $I : X \rightarrow [0, \infty]$ if

1. For each $s \geq 0$, the set $\Phi(s) \stackrel{\text{def}}{=} \{x \in X : I(x) \leq s\}$ is a compact subset of X .
2. For every open $G \subset X$,

$$\varliminf_{n \nearrow \infty} \frac{1}{n} \ln \mathbb{P} \{ \xi_n \in G \} \geq - \inf_{x \in G} I(x).$$

3. For every closed $F \subset X$,

$$\varlimsup_{n \nearrow \infty} \frac{1}{n} \ln \mathbb{P} \{ \xi_n \in F \} \leq - \inf_{x \in F} I(x).$$

One thing which is clear from [Example 3.2](#) is that we need to keep track of the type associated with each default (but not the types associated with names which do not default). To organize this,

let $\mathcal{M}_1(X)$ be the collection of measures ν on $(X, \mathcal{B}(X))$ such that $\nu(X) \leq 1$ (i.e., the collection of subprobability measures).

As it is discussed in Section 6, $\mathcal{M}_1(X)$ is a Polish space. For each $N \in \mathbb{N}$, define a random element ν_N of $\mathcal{M}_1(X)$ as

$$\nu_N \stackrel{\text{def}}{=} \frac{1}{N} \sum_{n=1}^N \Delta_n \delta_{\gamma^{N,n}}. \quad (4.1)$$

Thus, ν_N is the empirical measure on type-space determined by the names which default.

Next, define a sequence $(Z_N)_{N \in \mathbb{N}}$ of $[0, 1] \times \mathcal{M}_1(X)$ -valued random variables as

$$Z_N \stackrel{\text{def}}{=} (L_N, \nu_N) \quad N \in \mathbb{N}.$$

Since $[0, 1]$ and $\mathcal{M}_1(X)$ are both Polish, $[0, 1] \times \mathcal{M}_1(X)$ is also Polish. We seek a large deviations principle for the Z_N 's. Since projection maps are continuous, the contraction principle will then imply a large deviations principle for the L_N 's. Note furthermore that

$$D_N = \nu_N(X);$$

the map $\nu \mapsto \nu(X)$ is continuous in the topology on $\mathcal{M}_1(X)$, so the recovery statistics depend continuously on ν_N .

Before proceeding with the analysis, we need to define some more quantities.

Definition 4.2. For $p \in [0, 1]$, define

$$\tilde{h}_p(x) \stackrel{\text{def}}{=} \begin{cases} x \ln \frac{x}{p} + (1-x) \ln \frac{1-x}{1-p} & \text{for } x \text{ and } p \text{ in } (0, 1) \\ \ln \frac{1}{p} & \text{for } x = 1, p \in (0, 1] \\ \ln \frac{1}{1-p} & \text{for } x = 0, p \in [0, 1) \\ \infty & \text{else.} \end{cases}$$

For $\nu \in \mathcal{M}_1(X)$, define

$$H(\nu) \stackrel{\text{def}}{=} \begin{cases} \int_{\gamma=(p,q) \in X} \tilde{h}_p \left(\frac{d\nu}{dU}(\gamma) \right) U(d\gamma) & \text{if } \nu \ll U \\ \infty & \text{else.} \quad \square \end{cases}$$

To understand the behavior of L_N , we need to construct its moment generating function. For this purpose, we have the following definition.

Definition 4.3. For $\nu \in \mathcal{M}_1(X)$, define

$$\begin{aligned} \Lambda_\nu(\theta) &\stackrel{\text{def}}{=} \int_{\gamma=(p,q) \in X} \left\{ \ln \int_{r \in [0,1]} e^{\theta(1-r)} Q(\nu(X), dr) \right\} \nu(d\gamma) \quad \theta \in \mathbb{R} \\ \Lambda_\nu^*(\ell) &\stackrel{\text{def}}{=} \sup_{\theta \in \mathbb{R}} \{ \theta \ell - \Lambda_\nu(\theta) \} \quad \ell \in [0, 1] \end{aligned}$$

$\Lambda_\nu^*(\ell)$ is the Legendre–Fenchel transform of $\Lambda_\nu(\theta)$. \square

Let us now see if we can formally identify a large deviations principle for the Z_N 's in the setting of [Examples 3.1](#) and [3.2](#). In both examples we want to find a map $I : [0, 1] \times \mathcal{M}_1(X) \rightarrow [0, \infty]$ such that, at least formally,

$$\mathbb{P}_N \{Z_N \in dz^*\} \asymp \exp[-NI(z^*)] \quad N \rightarrow \infty$$

for each fixed $z^* = (\ell^*, \nu^*) \in [0, 1] \times \mathcal{M}_1(X)$. Assume that $\nu^* \ll U$. This is done without loss of generality, since, as we shall also see in the sequel, the rate function $I(z^*) = \infty$ if $\nu^* \not\ll U$.

Conditioning on the event that $\nu_N \approx \nu^*$, we have

$$\mathbb{P}_N \{Z_N \in dz^*\} = \mathbb{P} \{L_N \in d\ell^* | \nu_N \approx \nu^*\} \mathbb{P}_N \{\nu_N \in d\nu^*\}. \quad (4.2)$$

We want to derive asymptotic expressions for the first and second terms in the right-hand side of (4.2).

Example 4.4 (*Homogeneous Example 3.1 Continued*). First, we understand the second term in (4.2), i.e., $\mathbb{P}_N \{\nu_N \in d\nu^*\}$. Recall that

$$\nu_N(X) = \frac{1}{N} \sum_{1 \leq n \leq N} \Delta_n.$$

This is now essentially the focus of Sanov's theorem—i.i.d. Bernoulli random variables. Namely, recalling the definition of $\bar{h}_p(x)$ and $H(\nu)$ by [Definition 4.2](#) we have

$$\mathbb{P}_N \{\nu_N \in d\nu^*\} \asymp \exp[-N \bar{h}_p(\nu^*(X))] = \exp[-NH(\nu^*(X))]. \quad (4.3)$$

Second, we understand the first term in (4.2), i.e. $\mathbb{P} \{L_N \in d\ell^* | \nu_N \approx \nu^*\}$. As in [Example 3.2](#), we should have in law

$$L_N \approx \frac{1}{N} \left\{ \sum_{1 \leq n \leq N\nu^*(X)} (1 - \xi_n) \right\} \quad (4.4)$$

where the ξ_n 's are i.i.d. with common law $\varrho(\nu(X), \cdot)$. In the homogeneous pool of this example, the log moment generating function defined in [Definition 4.3](#) takes the form

$$\Lambda_{\nu^*}(\theta) = \nu^*(X) \ln \int_{r \in [0,1]} e^{\theta(1-r)} \varrho(\nu(X), dr)$$

for all $\theta \in \mathbb{R}$. The logarithmic moment generating function of L_N of (4.4) is approximately $N\Lambda_{\nu^*}(\theta/N)$. Thus we should have that

$$\mathbb{P}_N \{L_N \in d\ell^* | \nu_N \approx \nu^*\} \asymp \exp[-N\Lambda_{\nu^*}(\ell)].$$

We should then get the large deviations principle for Z_N by combining this, (4.3) and (4.2), i.e.,

$$\mathbb{P}_N \{Z_N \in dz^*\} \asymp \exp[-N[\Lambda_{\nu^*}(\ell) + H(\nu^*)]]. \quad \square$$

Example 4.5 (*Heterogeneous Example 3.2 Continued*). We proceed similarly to [Example 4.4](#). First, we understand the second term in (4.2), i.e., $\mathbb{P}_N \{\nu_N \in d\nu^*\}$. Observe, that we explicitly have

$$\frac{d\nu^*}{dU}(\gamma_A) = \frac{\nu^*\{\gamma_A\}}{1/3} \quad \text{and} \quad \frac{d\nu^*}{dU}(\gamma_B) = \frac{\nu^*\{\gamma_B\}}{2/3}.$$

Similarly

$$\frac{dv_N}{dU}(\gamma_A) = \frac{v_N\{\gamma_A\}}{U\{\gamma_A\}} = \frac{1}{N/3} \sum_{\substack{1 \leq n \leq N \\ n \in 3\mathbb{N}}} \Delta_n$$

$$\frac{dv_N}{dU}(\gamma_B) = \frac{v_N\{\gamma_B\}}{U\{\gamma_B\}} = \frac{1}{2N/3} \sum_{\substack{1 \leq n \leq N \\ n \notin 3\mathbb{N}}} \Delta_n.$$

Thus

$$\begin{aligned} \mathbb{P}_N\{v_N \in dv^*\} &= \mathbb{P}_N\{v_N\{\gamma_A\} \approx v^*\{\gamma_A\}, v_N\{\gamma_B\} \approx v^*\{\gamma_B\}\} \\ &= \mathbb{P}_N\left\{\frac{dv_N}{dU}(\gamma_A) \approx \frac{dv^*}{dU}(\gamma_A), \frac{dv_N}{dU}(\gamma_B) \approx \frac{dv^*}{dU}(\gamma_B)\right\}. \end{aligned}$$

Again, this is essentially the focus of Sanov's theorem—i.i.d. Bernoulli random variables. Namely, recalling the definition of $\hbar_p(x)$ and $H(v)$ by Definition 4.2 we have

$$\begin{aligned} \mathbb{P}_N\{v_N \in dv^*\} &\asymp \exp\left[-\frac{N}{3}\hbar_{p_A}\left(\frac{dv^*}{dU}(\gamma_A)\right) - \frac{2N}{3}\hbar_{p_B}\left(\frac{dv^*}{dU}(\gamma_B)\right)\right] \\ &= \exp\left[-N \int_{\gamma=(p,q) \in X} \hbar_p\left(\frac{dv^*}{dU}(\gamma)\right) U(d\gamma)\right] = \exp[-NH(v^*)]. \quad (4.5) \end{aligned}$$

Second, we understand the first term in (4.2), i.e. $\mathbb{P}\{L_N \in d\ell^* | v_N \approx v^*\}$. If $v_N \approx v^*$, then there are about $Nv^*\{\gamma_A\}$ defaults of type A, and $Nv^*\{\gamma_B\}$ defaults of type B. Thus, as in Example 3.2, we should have in law

$$L_N \approx \frac{1}{N} \left\{ \sum_{1 \leq n \leq Nv^*\{\gamma_A\}} (1 - \xi_n^A) + \sum_{1 \leq n \leq Nv^*\{\gamma_B\}} (1 - \xi_n^B) \right\} \quad (4.6)$$

where the ξ_n^A 's are i.i.d. with common law $\varrho_A(v(X), \cdot)$ and the ξ_n^B 's are i.i.d. with common law $\varrho_B(v(X), \cdot)$. In the heterogeneous pool of this example, the log moment generating function defined in Definition 4.3 takes the form

$$A_{v^*}(\theta) = v^*\{\gamma_A\} \ln \int_{r \in [0,1]} e^{\theta(1-r)} \varrho_A(v(X), dr) + v^*\{\gamma_B\} \ln \int_{r \in [0,1]} e^{\theta(1-r)} \varrho_B(v(X), dr)$$

for all $\theta \in \mathbb{R}$. The logarithmic moment generating function of L_N of (4.6) is approximately $NA_{v^*}(\theta/N)$. Thus we should have that

$$\mathbb{P}_N\{L_N \in d\ell^* | v_N \approx v^*\} \asymp \exp[-NA_{v^*}(\ell)].$$

We should then get the large deviations principle for Z_N by combining this, (4.5) and (4.2), i.e.,

$$\mathbb{P}_N\{Z_N \in dz^*\} \asymp \exp[-N[A_{v^*}^*(\ell) + H(v^*)]]. \quad \square$$

Examples 4.4 and 4.5 motivate our main result.

Theorem 4.6 (Main). *Let Assumption 3.3 hold and recall Definitions 4.2 and 4.3. Then, we have that $(Z_N)_{N \in \mathbb{N}}$ has a large deviations principle with rate function*

$$I(z) = A_v^*(\ell) + H(v)$$

for all $z = (\ell, \nu) \in [0, 1] \times \mathcal{M}_1(\mathbf{X})$. Second $(L_N)_{N \in \mathbb{N}}$ has a large deviations principle with rate function

$$I'(\ell) = \inf_{\nu \in \mathcal{M}_1(\mathbf{X})} I(\ell, \nu) \quad (4.7)$$

for all $\ell \in [0, 1]$. \square

Proof. Combine together [Propositions 6.1, 7.3](#) and [8.1](#). This gives the large deviations principle for (Z_N) . The large deviations principle for $(L_N)_{N \in \mathbb{N}}$ follows from the contraction principle and the continuity of the map $\nu \mapsto \nu(\mathbf{X})$. \square

Next we mention a useful representation formula for $H(\nu)$ that is defined in [Definition 4.2](#). Its proof will be given in [Section 10](#). Define

$$\lambda_p(\theta) \stackrel{\text{def}}{=} \ln(pe^\theta + 1 - p)$$

for all $\theta \in \mathbb{R}$ and $p \in [0, 1]$. Then λ_p and \hbar_p are convex duals; i.e.,

$$\begin{aligned} \hbar_p(x) &= \sup_{\theta \in \mathbb{R}} \{\theta x - \lambda_p(\theta)\} \quad x \in \mathbb{R} \\ \lambda_p(\theta) &= \sup_{x \in \mathbb{R}} \{\theta x - \hbar_p(x)\} \quad \theta \in \mathbb{R} \end{aligned} \quad (4.8)$$

and we have the following result.

Lemma 4.7. *Let [Assumption 3.3](#) hold. Then, we have that*

$$H(\nu) = \sup_{\phi \in C(\mathbf{X})} \left\{ \int_{\gamma \in \mathbf{X}} \phi(\gamma) \nu(d\gamma) - \int_{\gamma=(p,q) \in \mathbf{X}} \lambda_p(\phi(\gamma)) \mathbf{U}(d\gamma) \right\} \quad (4.9)$$

for all $\nu \in \mathcal{M}_1(\mathbf{X})$. \square

One way to interpret [Theorem 4.6](#) is that the rate functions I and I' give the correct way to find the “minimum-energy” configurations for atypically large losses to occur. In general, variational problems involving measures can be computationally difficult, so [Section 9](#) addresses some computational issues. In particular, we find an alternate expression which takes advantage of the specific structure of our problem. Define

$$\begin{aligned} M_\varrho(\theta, D) &\stackrel{\text{def}}{=} \ln \int_{r \in [0,1]} e^{\theta(1-r)} \varrho(D, dr) \quad \theta \in \mathbb{R} \\ I_\varrho(x, D) &\stackrel{\text{def}}{=} \sup_{\theta \in \mathbb{R}} \{\theta x - M_\varrho(\theta, D)\} \quad x \in \mathbb{R} \end{aligned} \quad (4.10)$$

for all $D \in [0, 1]$. Observe that we can write

$$\Lambda_\nu(\theta) = \int_{\gamma=(p,q) \in \mathbf{X}} M_\varrho(\theta, \nu(\mathbf{X})) \nu(d\gamma) \quad (4.11)$$

for all $\nu \in \mathcal{M}_1(\mathbf{X})$ and $\theta \in \mathbb{R}$. Define next $\mathcal{B} \stackrel{\text{def}}{=} B(\mathbf{X}; [0, 1])$.

Theorem 4.8. Let [Assumption 3.3](#) hold. For $\ell \in [0, 1]$ and $\mathbf{U} \in \mathcal{P}(\mathbf{X})$, set

$$J'(\ell) \stackrel{\text{def}}{=} \inf_{\Phi, \Psi \in \mathcal{B}} \left\{ \int_{\gamma=(p,q) \in \mathbf{X}} \left\{ \Phi(\gamma) I_{\varrho} \left(\Psi(\gamma), \int_{\gamma \in \mathbf{X}} \Phi(\gamma) \mathbf{U}(d\gamma) \right) + \hbar_p(\Phi(\gamma)) \right\} \mathbf{U}(d\gamma) : \right. \\ \left. \int_{\gamma \in \mathbf{X}} \Phi(\gamma) \Psi(\gamma) \mathbf{U}(d\gamma) = \ell \right\}. \quad (4.12)$$

We have that $I'(\ell) = J'(\ell)$ for all $\ell \in [0, 1]$. An alternate representation of J' is

$$J''(\ell) \stackrel{\text{def}}{=} \inf_{D \in [\ell, 1]} \inf_{\Phi \in \mathcal{B}} \inf_{\Psi \in \mathcal{B}} \left\{ \int_{\gamma=(p,q) \in \mathbf{X}} \left\{ \Phi(\gamma) I_{\varrho}(\Psi(\gamma), D) + \hbar_p(\Phi(\gamma)) \right\} \mathbf{U}(d\gamma) : \right. \\ \left. \int_{\gamma \in \mathbf{X}} \Phi(\gamma) \Psi(\gamma) \mathbf{U}(d\gamma) = \ell, \int_{\gamma \in \mathbf{X}} \Phi(\gamma) \mathbf{U}(d\gamma) = D \right\}. \quad \square \quad (4.13)$$

The proof of this is given in [Section 9](#). The point of the second representation [\(4.13\)](#) is that the innermost minimization problem (the one with Φ and D fixed) involves linear constraints. Namely, that Φ takes values in $[0, 1]$ and that two integrals of Φ take specific values. This will be useful in some of our numerical studies in the next section.

5. Examples and discussion

In [Section 5.1](#) we present some numerical examples. These examples showcase some of the possibilities and some of the implications of the dependence of the recovery distribution on the defaults. We conclude this section with [Section 5.2](#), where we summarize our conclusions.

5.1. Numerical examples

Let us see what our calculations look like in some specific cases. To focus on the effects of recovery, let us assume a common (relatively high) probability of default of 8%; i.e., $p^{N,n} = 0.08$ for all $N \in \mathbb{N}$ and $n \in \{1, 2, \dots, N\}$.

In the first group of examples (Cases 1–4), we will consider four specific cases, one with fixed recovery rate, two homogeneous pools with stochastic recoveries and one heterogeneous pool with stochastic recovery. For comparison purposes, the law of large numbers average loss in the pool will be the same in all cases and equal to $\bar{L} = 0.064$. We would like to remark the following. In a more realistic scenario one would expect to have pools with combinations of high-rated and low-rated counter-parties that have different default probabilities. Our formulation allows for such a scenario, but since we want to focus on the effects of recovery we assume a common probability of default in all of our examples. Moreover, recall that we have assumed as a reference point that all bonds have par value of \$1. We will see that the tails (the large deviations principle rate functions) are significantly different. Although our theory has primarily focused on the rate function in the large deviations principle for $(L_N)_{N \in \mathbb{N}}$, the solution of the variational problem [\(4.7\)](#) (or equivalently [\(4.12\)](#) or [\(4.13\)](#)) gives useful information.

In the second group of examples (Cases 5–6), we consider two heterogeneous pools, one with recovery whose distribution depends on the default rate (Case 5) and one with recovery whose distribution does not depend on the default rate (Case 6). As expected, the dependence on the default rate affects the tails.

In Cases 2–5 we consider stochastic recovery rates. We want to consider the case that the recovery is in an appropriate sense negatively correlated with the defaults; i.e., that more defaults

imply less recovery. In an economy that experiences recession, recovery rates tend to decrease just as defaults tend to increase. This property is also a documented empirical observation, e.g., see [15,1] and the references therein.

In *Case 1*, let us assume that the recovery rate is fixed at 20%; i.e., $q^{N,n} = \delta_{0.2}$ for all $N \in \mathbb{N}$ and $n \in \{1, 2, \dots, N\}$. The LDP here is essentially a straightforward application of Cramér's theorem. Let us see how our general formulation of [Theorem 4.6](#) covers this as a special case. In this case $\gamma = (0.08, \delta_{0.2})$ and $\mathbf{U} = \delta_\gamma$. Also we have

$$H(v) = \begin{cases} \hbar_{0.08}(v(\mathbf{X})) & \text{if } v = v(\mathbf{X})\delta_\gamma \\ \infty & \text{else} \end{cases}$$

and

$$\Lambda_v(\theta) = 0.8\theta v(\mathbf{X}) \quad \text{and} \quad \Lambda_v^*(\ell) = \begin{cases} 0 & \text{if } \ell = 0.8v(\mathbf{X}) \\ \infty & \text{else.} \end{cases}$$

Collecting things together, we have that $(Z_N)_{N \in \mathbb{N}}$ and $(L_N)_{N \in \mathbb{N}}$ are governed, respectively, by the rate functions

$$I_1(\ell, v) = \begin{cases} \hbar_{0.08}\left(\frac{\ell}{0.8}\right) & \text{if } v = \frac{\ell}{0.8}\delta_\gamma \\ \infty & \text{else} \end{cases}$$

$$I_1'(\ell) = \hbar_{0.08}\left(\frac{\ell}{0.8}\right).$$

We note that $I_1'(\ell)$ is finite only if $0 \leq \ell \leq 0.8$.

In *Case 2* we consider a homogeneous pool with the recovery rate following a beta distribution. Essentially, this is a special case of [Examples 3.1](#) and [4.4](#). For $\beta > 0$, define

$$\mu_\beta(A) \stackrel{\text{def}}{=} \beta \int_{r \in A} (1-r)^{\beta-1} dr, \quad A \in \mathcal{B}[0, 1]$$

this is the law of the beta distribution with parameters 1 and β . As β increases, the amount of mass near 1 decreases. We also have that

$$\int_{r \in [0, 1]} r \mu_\beta(dr) = \frac{1}{1+\beta}$$

for all $\beta > 0$ (as β increases, the mean of μ_β decreases). This will allow a number of explicit formulas for the expected recovery (given the default rate).

We here assume that the recovery rate has a beta distribution whose parameters depend linearly and monotonically on the empirical default rate. Namely, if the default rate is D , then the recoveries will all have common beta distribution with parameters 1 and

$$\beta = f_{\text{aff}}(D) \stackrel{\text{def}}{=} \frac{1}{0.2 - 0.1(D - 0.08)} - 1;$$

note that $f : [0, 1] \rightarrow \mathbb{R}_+$. Define

$$q_{\text{aff}}(D, \cdot) = \mu_{f_{\text{aff}}(D)} \quad \text{for all } D \in [0, 1]$$

$$\gamma_{\text{aff}} = (0.08, q_{\text{aff}}).$$

Then $\mathbf{U} = \delta_{\gamma_{\text{aff}}}$.

This choice of f_{aff} results in a conditional expected recovery which is affine in D ; i.e.,

$$\int_{r \in [0,1]} r \varrho_{\text{aff}}(D, dr) = 0.2 - 0.1(D - 0.08).$$

Observe that the expected recovery of a single name in the homogeneous pool is decreasing in D and always in $[0, 1]$. According to (3.1), the law of large numbers average loss is

$$\bar{L} = 0.08 \times \left(1 - \frac{1}{1 + f_{\text{aff}}(0.08)} \right) = 0.08 \times (1 - 0.2) = 0.064.$$

The rate functions in this case are somewhat similar to those in Case 1. Again we have that

$$H(v) = \begin{cases} \check{h}_{0.08}(v(\mathbf{X})) & \text{if } v = v(\mathbf{X})\delta_{\gamma_{\text{aff}}} \\ \infty & \text{else.} \end{cases}$$

For each $\beta > 0$, define

$$\check{A}_{\beta}(\theta) \stackrel{\text{def}}{=} \ln \int_{r \in [0,1]} e^{\theta(1-r)} \mu_{\beta}(dr) \quad \theta \in \mathbb{R}$$

$$\check{A}_{\beta}^*(\ell) \stackrel{\text{def}}{=} \sup_{\theta \in \mathbb{R}} \left\{ \theta \ell - \check{A}_{\beta}(\theta) \right\} \quad \ell \in \mathbb{R}.$$

Then if $v = v(\mathbf{X})\delta_{\gamma_{\text{aff}}}$ where $v(\mathbf{X}) > 0$ we have from Definition 4.3,

$$A_v(\theta) = v(\mathbf{X})\check{A}_{f_{\text{aff}}(v(\mathbf{X}))}(\theta) \quad \theta \in \mathbb{R}$$

$$A_v^*(\ell) = v(\mathbf{X})\check{A}_{f_{\text{aff}}(v(\mathbf{X}))}^*\left(\frac{\ell}{v(\mathbf{X})}\right) \quad \ell \in \mathbb{R}.$$

Here $(Z_N)_{N \in \mathbb{N}}$ and $(L_N)_{N \in \mathbb{N}}$ are governed, respectively, by the rate functions.

$$I_2(\ell, v) = \begin{cases} \check{h}_{0.08}(v(\mathbf{X})) + v(\mathbf{X})\check{A}_{f_{\text{aff}}(v(\mathbf{X}))}^*\left(\frac{\ell}{v(\mathbf{X})}\right) & \text{if } v = v(\mathbf{X})\delta_{\gamma_{\text{aff}}} \\ \infty & \text{else} \end{cases} \quad (5.1)$$

$$I_2'(\ell) = \inf_{D \in (0,1]} \left\{ \check{h}_{0.08}(D) + D\check{A}_{f_{\text{aff}}(D)}^*\left(\frac{\ell}{D}\right) \right\}.$$

In Case 3, we replace f_{aff} of Case 2 with one that results in a conditional expected recovery which is quadratic in D ; this allows us some insight into the effects of convexity in the conditional expected recovery. We set

$$f_q(D) = \frac{1}{0.2 - 0.1(D - 0.08) - 0.1(D - 0.08)^2} - 1.$$

Again we have that $f_q : [0, 1] \rightarrow \mathbb{R}_+$, and we set $\varrho_q(D, \cdot) = \mu_{f_q(D)}$, $\gamma_q = (0.08, \varrho_q)$, and have that $\mathbf{U} = \delta_{\gamma_q}$. Here we have that

$$\int_{r \in [0,1]} r \varrho_q(D, dr) = 0.2 - 0.1(D - 0.08) - 0.1(D - 0.08)^2.$$

Observe that the expected recovery of a single name in the homogeneous pool is decreasing in D and always in $[0, 1]$.

We again get that $\bar{L} = 0.064$. The corresponding rate function is I_3' .

Case 4 involves the beta distribution again. Here, however, we now consider a heterogeneous pool of two types (Examples 3.2 and 4.5). We concentrate on the effect of the heterogeneity in the recovery distribution, so as in the previous cases all bonds will have default probability of 8%. For all $D \in [0, 1]$, every third bond will have recovery distribution governed by ϱ_{aff} and the remaining bonds will have recovery distribution governed by ϱ_q . We thus have that $\mathbf{U} = \frac{1}{3}\delta_{\gamma_{\text{aff}}} + \frac{2}{3}\delta_{\gamma_q}$. It is easy to see that the average loss is

$$\bar{L} = 0.08 \times \frac{1}{3} \times \left(1 - \frac{1}{1 + f_{\text{aff}}(0.08)}\right) + 0.08 \times \frac{2}{3} \times \left(1 - \frac{1}{1 + f_q(0.08)}\right) = 0.064.$$

For notational convenience and in order to illustrate the usage of Theorem 4.8 we use the alternative representation (4.13). If $\varrho(D, \cdot) = \mu_{f(D)}$ for some $f \in C([0, 1]; \mathbb{R}_+)$, then for all $D \in [0, 1]$, we have that $M_\varrho(\theta, D) = \check{\Lambda}_{f(D)}(\theta)$ for all $\theta \in \mathbb{R}$ and $I_\varrho(\ell, D) = \check{\Lambda}_{f(D)}^*(\ell)$ for all $\ell \in [0, 1]$. Since the support of \mathbf{U} is exactly $\{\gamma_{\text{aff}}, \gamma_q\}$, we can consider Φ and Ψ in $\mathcal{B} = B(X; [0, 1])$ of the form

$$\Phi = \phi_A \chi_{\{\gamma_{\text{aff}}\}} + \phi_B \chi_{\{\gamma_q\}} \quad \text{and} \quad \Psi = \psi_A \chi_{\{\gamma_{\text{aff}}\}} + \psi_B \chi_{\{\gamma_q\}}.$$

Thus (4.13) becomes

$$\begin{aligned} I'_4(\ell) &= \inf_{D \in [\ell, 1]} \inf_{\psi_A, \psi_B \in [0, 1]} \inf_{\substack{\phi_A, \phi_B \in [0, 1] \\ \phi_A \psi_A / 3 + 2\phi_A \psi_B / 3 = \ell \\ \phi_A / 3 + 2\phi_B / 3 = D}} \left\{ \frac{1}{3} \phi_A I_{\varrho_{\text{aff}}}(\psi_A, D) + \frac{2}{3} \phi_B I_{\varrho_q}(\psi_B, D) \right. \\ &\quad \left. + \frac{1}{3} h_{p_A}(\phi_A) + \frac{2}{3} h_{p_B}(\phi_B) \right\} \\ &= \inf_{D \in [\ell, 1]} \inf_{\psi_A, \psi_B \in [0, 1]} \inf_{\substack{\phi_A, \phi_B \in [0, 1] \\ \phi_A \psi_A / 3 + 2\phi_A \psi_B / 3 = \ell \\ \phi_A / 3 + 2\phi_B / 3 = D}} \left\{ \frac{1}{3} \phi_A \check{\Lambda}_{f_{\text{aff}}(D)}^*(\psi_A) + \frac{2}{3} \phi_B \check{\Lambda}_{f_q(D)}^*(\psi_B) \right. \\ &\quad \left. + \frac{1}{3} h_{p_A}(\phi_A) + \frac{2}{3} h_{p_B}(\phi_B) \right\}. \end{aligned} \quad (5.2)$$

In Fig. 1, we plot the rate functions I'_1 , I'_2 , I'_3 and I'_4 . We use a Monte Carlo procedure to compute $\check{\Lambda}$ and $\check{\Lambda}^*$. As expected, all action functions are nonnegative and zero at the (common) law of large numbers average loss of $\bar{L} = 0.064$. We observe that $I'_3 \leq I'_4 \leq I'_2 \leq I'_1$. In particular, the heterogeneous case, which is a mixture of an affine conditional expected recovery and a quadratic conditional expected recovery, is in between the two homogeneous cases (Cases 2 and 3). Of course, we should not be surprised that the rate function in Case 1 is larger than that in Cases 2 through 4, there are in general many more configurations which lead to a given overall loss rate. An interesting observation that seems to be suggested by the examples considered here is that more convexity results in smaller values for the rate function.

The discussion from now on will be formal. However, we believe that it offers some useful insights into the effect of default rates on the average recovery of large pools.

Another useful insight which we can numerically extract is the “preferred” way which losses stem from defaults versus recovery. For each $\ell \in [0, 1]$, let $D_*(\ell)$ be the minimizer² in the

² Case 1 is of course degenerate in this sense; for a given loss rate ℓ , the default rate must be very close to $\ell/0.8$.

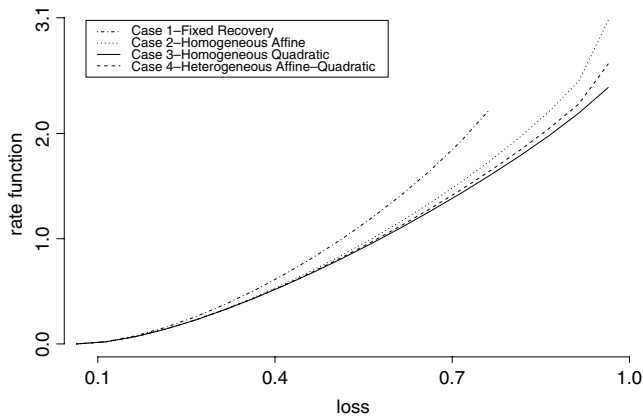


Fig. 1. Rate functions for fixed recovery, for the homogeneous and heterogeneous cases.

expression (5.1) for I'_2, I'_3 or alternately the expression (5.2) for I'_4 . $D_*(\ell)$ can be interpreted as the “most likely” default rate in the pool given that $L_N \approx \ell$. We make the following assumption.

Assumption 5.1. Assume that for each $\ell \in [0, 1]$ the aforementioned minimizer $D_*(\ell)$ exists and is unique.

Numerically finding $D_*(\ell)$ in the examples considered here, verifies that Assumption 5.1 holds in these cases. For $\ell > \bar{L}$ and $\delta > 0$, the Gibbs conditioning principle [5, Section 7.3] implies that we should have that

$$\lim_{N \rightarrow \infty} \mathbb{P}_N \{ |D_N - D_*(\ell)| \geq \delta | L_N \geq \ell \} = 0. \quad (5.3)$$

In other words, conditional on the pool suffering losses exceeding rate ℓ , the default rate should converge to $D_*(\ell)$. Using this information, we can then say something about how the average recovery of the pool is related to the optimal default rate $D_*(\ell)$. Motivated by (3.1) we write that

$$\text{Average Loss} = \text{Default} \times (1 - \text{Recovery}) \Rightarrow \text{Recovery} = 1 - \frac{\text{Average Loss}}{\text{Default}}$$

to find an effective recovery rate in terms of the loss rate and the default rate. This formulation quantifies the fact that losses are due to *both* default and recovery. For atypically large losses in a large pool of credit assets, we should combine this with the Gibbs conditioning calculation of (5.3). Namely, let us define

$$\mathcal{R}(\ell) = 1 - \frac{\ell}{D_*(\ell)}.$$

We can also formalize the dependence of recovery on default by letting $\mathcal{R}^* : [0, 1] \rightarrow [0, 1]$ be such that

$$\mathcal{R}^*(D_*(\ell)) = \mathcal{R}(\ell) \quad \text{for all } \ell \in [0, 1]. \quad (5.4)$$

We refer to $\mathcal{R}^*(\cdot)$ as the *effective average recovery* of the pool.

Fig. 2 is a plot of \mathcal{R}^* for the cases which we are studying. As it was expected, we observe that, for Cases 2–4, the optimal average recovery of the pool and the optimal defaults are negatively correlated. Of course, this is consistent with the corresponding negative correlation

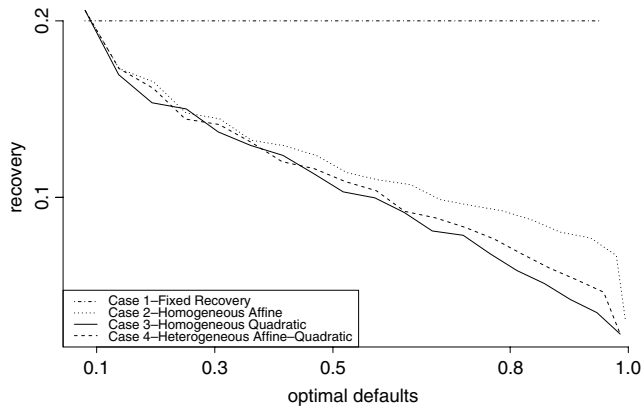


Fig. 2. Effective average recovery of the pool versus optimal defaults.

of the individual recovery and default rate that is embedded in the choices of the individual recovery distribution. Notice however that \mathcal{R}^* is a “global” quantity in that it represents the effective average recovery in the pool. Another interesting observation is that the graph of the heterogeneous case which is a mixture of an affine and a quadratic conditional expected recovery is between the graph of the homogeneous Cases 2 and 3 which treat the affine and quadratic cases respectively.

We conclude this subsection with a comparison of a heterogeneous portfolio whose distribution of recovery depends on the default rate with one that does not.

Case 5. Consider the setup of Case 4 with the only difference being the definitions of $f_{\text{aff}}(D)$ and $f_q(D)$. In particular, we consider

$$f_{\text{aff}}(D) = \frac{1}{0.1 - 0.05(D - 0.08)} - 1,$$

$$f_q(D) = \frac{1}{0.25 - 0.1(D - 0.08) - 0.1(D - 0.08)^2} - 1.$$

The corresponding rate function is I'_5 .

Case 6. Again, consider the setup of Case 4 with the only difference being the definitions of $f_{\text{aff}}(D)$ and $f_q(D)$. In particular, we consider

$$f_{\text{aff}} = \frac{1}{0.1} - 1,$$

$$f_q = \frac{1}{0.25} - 1.$$

Notice that in this case the distribution of the recovery is independent of the default rate D . The corresponding rate function is I'_6 .

In both cases the law of large numbers average loss is the same as before. In Fig. 3, we plot (a) on the left, the rate functions I'_5 and I'_6 and (b) on the right, \mathcal{R}^* for Cases 5 and 6.

As it is indicated by the left figure, one of the effects of the dependence of the distribution of the recovery on the defaults is to decrease the values of the rate function, i.e., $I'_5 \leq I'_6$. The effective average recovery, \mathcal{R}^* , is in both cases negatively correlated with the optimal default rate D_* . Notice however that for large overall default rates the effective average recovery of the heterogeneous pool with dependence on defaults (Case 5) is less than the effective average

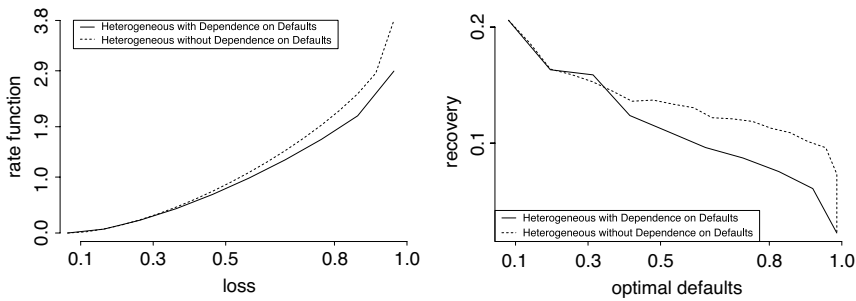


Fig. 3. Rate functions and \mathcal{R}^* when recovery distribution depends on defaults and when it does not.

recovery of the heterogeneous pool without dependence on defaults (Case 6). This is consistent with our intuition, namely that dependence of the recoveries on defaults should affect the recovery of the pool and in particular that more defaults should decrease recovery.

5.2. Conclusions and discussion

In this subsection we summarize our findings and pose some questions that would be interesting to study.

- Assuming that the recoveries for the defaulted assets depend upon the number of defaults in a fairly general way we have characterized the typical (Lemma 3.5) and atypical (Theorems 4.6 and 4.8) behavior of the loss rates in the pool. This allows for consideration of the case when recovery rates are affected by the number of defaults. We prove in Theorem 4.6 the large deviations principle for the joint family of random variables (L_N, ν_N) where L_N is the average loss and ν_N is the empirical measure on type-space determined by the names which default. Then, the large deviations principle for L_N follows by the contraction principle. Furthermore, we derive in Theorem 4.8 various equivalent representations for the rate function which give some more insight to the favored way to rearrange recoveries and losses among the different types and also ease its numerical computation.
- Moreover, as we have demonstrated in Section 5.1, the rate functions can be calculated numerically for both homogeneous and heterogeneous cases. The rate function determines the main term in the logarithmic asymptotics of the probability that the average loss in the pool is, let us say, bigger than a specific level. In particular, it determines the main term of the tail distribution of total losses on a portfolio consisting of many positions. See also [4,9] for a related discussion. Also note that in the examples considered here, more convexity resulted in smaller rate functions.
- Our formulation allows to extract some information regarding the optimal default rate in the pool for a given level of average loss in the pool. Moreover, a useful insight into the fact that losses are due to *both* default and recovery is perhaps given by the *effective average recovery* as defined by relation (5.4).
- As it is indicated by the comparison of Cases 5 and 6, the effect of the dependence of the distribution of the recovery on the default rate is (a) to reduce the rate function and (b) to reduce the effective average recovery of the pool especially when the optimal default rate gets larger.

Some interesting questions that naturally arise are given below. These questions will be pursued elsewhere.

- In this paper we deal with logarithmic asymptotics. It would be interesting to study the exact asymptotics and characterize the prefactor that appears in the asymptotic expansion of the loss distribution as the pool gets larger. Similar questions have also been studied in [4,11] under various assumptions.
- A question that is in particular relevant for financial applications is to study measures such as Value-at-Risk (VaR) and expected shortfall (ES). VaR at level $a \in (0, 1)$ is the a -quantile of the loss distribution and expected shortfall at level a is defined as $ES_a = \mathbb{E}[L | L \geq \text{VaR}_a(L)]$. It would be interesting to study the asymptotic behavior of $\text{VaR}_a(L_N)$ and $ES_a(L_N)$, characterize their limits as the pool gets larger and study the corresponding implications of the dependence of the recoveries on defaults. Under certain conditions similar questions have also been investigated in [10,6,7].

6. Compactness of level sets

The first part of the large deviations claim is that the level sets of I are compact. The proof follows along fairly standard lines.

First, however, we need to topologize $\mathcal{M}_1(\mathbf{X})$. This is done in the usual way. In particular, fix a point \star that is not in \mathbf{X} and define $\mathbf{X}^+ \stackrel{\text{def}}{=} \mathbf{X} \cup \{\star\}$. Give \mathbf{X}^+ the standard topology; open sets are those which are open subsets of \mathbf{X} (with its original topology) or complements in \mathbf{X}^+ of closed subsets of \mathbf{X} (again, in the original topology of \mathbf{X}). Define a bijection ι from $\mathcal{M}_1(\mathbf{X})$ to $\mathcal{P}(\mathbf{X}^+)$ by setting

$$(\iota\nu)(A) \stackrel{\text{def}}{=} \nu(A \cap \mathbf{X}) + (1 - \nu(\mathbf{X})) \delta_\star(A)$$

for all $A \in \mathcal{B}(\mathbf{X}^+)$. The point \star is introduced because ν is a subprobability measure. The topology of $\mathcal{M}_1(\mathbf{X})$ is the pullback of the topology of $\mathcal{P}(\mathbf{X}^+)$ and the metric on $\mathcal{M}_1(\mathbf{X})$ is that given by requiring ι to be an isometry.

Since \mathbf{X} is Polish, so is \mathbf{X}^+ , and thus $\mathcal{P}(\mathbf{X}^+)$ is Polish, and thus $\mathcal{M}_1(\mathbf{X})$ is a Polish space.

Proposition 6.1. *For each $s \geq 0$, the set*

$$\Phi(s) \stackrel{\text{def}}{=} \{z \in [0, 1] \times \mathcal{M}_1(\mathbf{X}) : I(z) \leq s\}$$

is a compact subset of $[0, 1] \times \mathcal{M}_1(\mathbf{X})$. \square

Proof. We first claim that $\Phi(s)$ is contained in a compact subset of $[0, 1] \times \mathcal{M}_1(\mathbf{X})$. Since $[0, 1]$ is already compact, it suffices to show that $\Phi_M(s) \stackrel{\text{def}}{=} \{\nu \in \mathcal{M}_1(\mathbf{X}) : H(\nu) \leq s\}$ is a compact subset of $\mathcal{M}_1(\mathbf{X})$. If $\nu \in \Phi_M(s)$, then $\nu \ll \mathbf{U}$ and, since $h_p(x) = \infty$ for $x > 1$, we have that

$$\mathbf{U} \left\{ \gamma \in \mathbf{X} : \frac{d\nu}{d\mathbf{U}}(\gamma) > 1 \right\} = 0,$$

so for any $A \in \mathcal{B}(\mathbf{X})$, $\nu(A) \leq \mathbf{U}(A)$. Since \mathbf{U} itself is tight (it is a probability measure on a Polish space), $\Phi_M(s)$ is tight; for every $\varepsilon > 0$, there is a $K_\varepsilon \subset\subset \mathbf{X}$ such that $\nu(\mathbf{X} \setminus K_\varepsilon) < \varepsilon$ for all $\nu \in \Phi_M(s)$. We claim that thus $\iota(\Phi_M(s))$ is also tight. Indeed, fix $\varepsilon > 0$. Letting $\iota_o : \mathbf{X} \rightarrow \mathbf{X}^+$ be the inclusion map, we have that ι_o is continuous, and thus $\iota_o(K_\varepsilon)$ is compact. Since singletons are also compact, $K^* \stackrel{\text{def}}{=} \iota_o(K_\varepsilon) \cup \{\star\}$ is a compact subset of \mathbf{X}^+ . For every $\nu \in \Phi_M(s)$, $(\iota\nu)(\mathbf{X}^+ \setminus K^*) = \nu(\mathbf{X} \setminus K_\varepsilon) < \varepsilon$, so indeed $\iota(\Phi_M(s))$ is tight. Thus $\overline{\Phi_M(s)} \subset\subset \mathcal{P}(\mathbf{X}^+)$ and hence

$$\Phi_M(s) \subset \iota^{-1}(\overline{\iota(\Phi_M(s))}) \subset\subset \mathcal{M}_1(\mathbf{X})$$

the last claim following since ι is a homeomorphism. Gathering things together, we have that $\Phi(s)$ is indeed contained in a compact subset of $[0, 1] \times \mathcal{M}_1(X)$.

We now want to show that $\Phi(s)$ is closed, or equivalently, that $([0, 1] \times \mathcal{M}_1(X)) \setminus \Phi(s)$ is open. Using Lemma 4.7, we have that

$$([0, 1] \times \mathcal{M}_1(X)) \setminus \Phi(s) = \bigcup_{\substack{\theta \in \mathbb{R} \\ \phi \in C(X)}} \left\{ (\ell, \nu) \in \mathcal{M}_1(X) : \theta\ell + \int_{\gamma \in X} \phi(\gamma) \nu(d\gamma) - \Lambda_\nu(\theta) > s \right. \\ \left. + \int_{\gamma=(p, wp) \in X} \lambda_p(\phi(\gamma)) U(d\gamma) \right\}.$$

For each $\theta \in \mathbb{R}$ and $\phi \in C(X)$, then map $(\ell, \nu) \mapsto \theta\ell + \int_{\gamma \in X} \phi(\gamma) \nu(d\gamma) - \Lambda_\nu(\theta)$ is continuous, so we have written $([0, 1] \times \mathcal{M}_1(X)) \setminus \Phi(s)$ as a union of open sets. \square

7. Large deviations lower bound

We next prove the large deviations lower bound. As with most large deviations lower bounds, the idea is to find a measure transformation under which the set of interest becomes “typical”. In this case, this measure transformation will come from a combination of Cramér’s theorem and Sanov’s theorem.

First, we mention an auxiliary approximation result which will be useful in the proof. Its proof is in Section 10.

Lemma 7.1. Fix $\nu \in \mathcal{M}_1(X)$ such that $H(\nu) < \infty$. Then there is a sequence $(\nu_N)_{N \in \mathbb{N}}$ such that

$$\lim_{N \rightarrow \infty} \nu_N = \nu \quad \text{and} \quad \lim_{N \rightarrow \infty} H(\nu_N) = H(\nu)$$

(and thus $\nu_N \ll U$ for all $N \in \mathbb{N}$) and such that

$$\gamma \mapsto \frac{d\nu_N}{dU}(\gamma) \quad \text{and} \quad \gamma = (p, \varrho) \mapsto \chi_{(0,1)}(p) h'_p \left(\frac{d\nu_N}{dU}(\gamma) \right)$$

are both well-defined and in $C(X)$ for all $N \in \mathbb{N}$.

We start with a simplified lower bound where the measure transformation in Cramér’s theorem is fairly explicit. For each $\nu \in \mathcal{M}_1(X)$, we make the usual definition [5, Appendix A] that

$$\text{dom} \Lambda_\nu^* \stackrel{\text{def}}{=} \{ \ell \in [0, 1] : \Lambda_\nu^*(\ell) < \infty \};$$

this will of course be an interval; $\text{ri dom} \Lambda_\nu^*$ will be the relative interior of $\text{dom} \Lambda_\nu^*$.

Proposition 7.2. Fix an open subset G of $[0, 1] \times \mathcal{M}_1(X)$ and $z = (\ell, \nu) \in G$ such that $I(z) < \infty$ and $\ell \in \text{ri dom} \Lambda_\nu^*$. Then

$$\lim_{N \rightarrow \infty} \frac{1}{N} \ln \mathbb{P}_N \{ Z_N \in G \} \geq -I(z). \quad \square \tag{7.1}$$

Proof. The proof will require a number of tools. For presentation purposes we split the proof in three steps. In Step 1, we prove several auxiliary results and are then used to define and analyze the measure change that is used to prove the initial lower bound. In Step 2, we prove that under the measure defined by (7.7) in Step 1, the recovery rates for the names that have defaulted are

independent and that the default probabilities are independent as well. In Step 3, we put things together in order to prove the initial lower bound (7.1).

Step 1. Since $\ell \in \text{ri dom } A_v^*$, there is a $\theta \in \mathbb{R}$ such that

$$A'_v(\theta) = \ell \quad \text{and} \quad A_v^*(\ell) = \theta A'_v(\theta) - A_v(\theta) \quad (7.2)$$

(see [5, Appendix A]). Let us now fix a relaxation parameter $\eta > 0$. Then there is an $\eta_1 \in (0, \eta)$ and an open neighborhood \mathcal{O}_1 of v such that $(\ell - \eta_1, \ell + \eta_1) \times \mathcal{O}_1 \subset G$. Using the first equality of (7.2), we have that $(A'_v(\theta), v) = (\ell, v) \in (\ell - \eta_1, \ell + \eta_1) \times \mathcal{O}_1$. Since the maps $(\tilde{\eta}, \tilde{v}) \rightarrow (A'_v(\theta) + \tilde{\eta}, \tilde{v})$ and $\tilde{v} \mapsto A_{\tilde{v}}(\theta)$ are continuous, there is an $\eta_2 \in (0, 1)$ and an open subset \mathcal{O}_2 of $\mathcal{M}_1(X)$ such that

$$\begin{aligned} \{ (A'_v(\theta) + \tilde{\eta}, \tilde{v}) : \tilde{\eta} \in (0, \eta_2), \tilde{v} \in \mathcal{O}_2 \} &\subset (\ell - \eta_1, \ell + \eta_1) \times \mathcal{O}_1 \\ |A_{\tilde{v}}(\theta) - A_v(\theta)| &< \eta \quad \text{for } \tilde{v} \in \mathcal{O}_2. \end{aligned} \quad (7.3)$$

We next want to use Lemma 7.1 to choose a particularly nice element of \mathcal{O}_2 . Namely, Lemma 7.1 ensures that there is a $v^* \in \mathcal{O}_2$ such that $v^* \ll \mathbf{U}$ and such that both $\frac{dv^*}{d\mathbf{U}}$ and

$$\phi(\gamma) \stackrel{\text{def}}{=} \chi_{(0,1)}(p) \hbar'_p \left(\frac{dv^*}{d\mathbf{U}}(\gamma) \right) \quad \gamma = (p, \varrho) \in X \quad (7.4)$$

are in $C(X)$ and such that $|H(v^*) - H(v)| < \eta$. Let \mathcal{O}_3 be an open subset of \mathcal{O}_2 which contains v^* and such that

$$\left| \int_{\gamma \in X} \phi(\gamma) \tilde{v}(d\gamma) - \int_{\gamma \in X} \phi(\gamma) v^*(d\gamma) \right| < \eta$$

for all $\tilde{v} \in \mathcal{O}_3$.

We can now proceed with our measure change. For each $N \in \mathbb{N}$, define

$$A_1^{(N)} \stackrel{\text{def}}{=} \theta L_N - A_{v_N}(\theta) \quad \text{and} \quad A_2^{(N)} \stackrel{\text{def}}{=} \frac{1}{N} \sum_{n=1}^N \Delta_n \phi(\gamma^{N,n}) - \lambda_{p^{N,n}}(\phi(\gamma^{N,n})). \quad (7.5)$$

Standard calculations show that

$$\mathbb{E}_N \left[\exp \left[N A_1^{(N)} \right] \middle| \mathcal{D} \right] = 1 \quad \text{and} \quad \mathbb{E}_N \left[\exp \left[N A_2^{(N)} \right] \right] = 1. \quad (7.6)$$

Define a new probability measure as

$$\tilde{\mathbb{P}}_N(A) \stackrel{\text{def}}{=} \mathbb{E}_N \left[\chi_A \exp \left[N \left\{ A_1^{(N)} + A_2^{(N)} \right\} \right] \right] \quad A \in \mathcal{B}(E). \quad (7.7)$$

This will be the desired measure change.

Define

$$S_N \stackrel{\text{def}}{=} \{ |L_N - A'_{v_N}(\theta)| < \eta_2, v_N \in \mathcal{O}_3 \}. \quad (7.8)$$

On S_N ,

$$(L_N, v_N) = (A'_{v_N}(\theta) + \{L_N - A'_{v_N}(\theta)\}, v_N) \in (\ell - \eta_1, \ell + \eta_1) \times \mathcal{O}_1 \subset G$$

so in fact $S_N \subset G$. Thus

$$\mathbb{P}_N\{Z_N \in G\} \geq \tilde{\mathbb{P}}_N \left[\chi_{S_N} \exp \left[-N \left\{ A_1^{(N)} + A_2^{(N)} \right\} \right] \right].$$

Let us also assume that N is large enough that

$$\left| \int_{\gamma=(p,\varrho)\in X} \lambda_p(\phi(\gamma)) \mathbf{U}_N(d\gamma) - \int_{\gamma=(p,\varrho)\in X} \lambda_p(\phi(\gamma)) \mathbf{U}(d\gamma) \right| < \eta.$$

Thus by (7.2), (7.3), (7.5) and (7.8) we have that

$$A_1^{(N)} \leq \theta \ell + |\theta| \eta_1 - \Lambda_v(\theta) + \eta = \Lambda_v^*(\ell) + (|\theta| + 1) \eta.$$

Next we prove that

$$\hbar_p \left(\frac{dv^*}{d\mathbf{U}}(\gamma) \right) = \frac{dv^*}{d\mathbf{U}}(\gamma) \phi(\gamma) - \lambda_p(\phi(\gamma)) \quad (7.9)$$

for \mathbf{U} -almost all $\gamma = (p, \varrho) \in X$. This follows from standard convex analysis and the form (7.4) of ϕ when $p \in (0, 1)$. Since $H(v^*) < \infty$, (10.2) implies that, except on a \mathbf{U} -negligible set,

$$\hbar_p \left(\frac{dv^*}{d\mathbf{U}}(\gamma) \right) = \hbar_p(p) = 0 = p \times 0 - \lambda_p(0) = \frac{dv^*}{d\mathbf{U}}(\gamma) \phi(\gamma) - \lambda_p(\phi(\gamma))$$

if $\gamma = (p, \varrho) \in X$ is such that $p \in \{0, 1\}$. In other words, (7.9) holds except on a \mathbf{U} -negligible set.

Therefore,

$$\begin{aligned} A_2^{(N)} &= \int_{\gamma \in X} \phi(\gamma) v_N(d\gamma) - \int_{\gamma=(p,\varrho)\in X} \lambda_p(\phi(\gamma)) \mathbf{U}_N(d\gamma) \\ &\leq \int_{\gamma \in X} \phi(\gamma) v^*(d\gamma) - \int_{\gamma=(p,\varrho)\in X} \lambda_p(\phi(\gamma)) \mathbf{U}(d\gamma) + 2\eta \\ &= \int_{\gamma=(p,\varrho)\in X} \left\{ \phi(\gamma) \frac{dv^*}{d\mathbf{U}}(\gamma) - \lambda_p(\phi(\gamma)) \right\} \mathbf{U}(d\gamma) + 2\eta \\ &= \int_{\gamma=(p,\varrho)\in X} \hbar_p \left(\frac{dv^*}{d\mathbf{U}}(\gamma) \right) \mathbf{U}(d\gamma) + 2\eta \leq H(v) + 3\eta. \end{aligned} \quad (7.10)$$

The last line of (7.10) follows from (7.9).

Thus

$$\mathbb{P}_N\{Z_N \in G\} \geq \tilde{\mathbb{P}}_N(S_N) \exp[-N\{I(z) - (|\theta| + 4)\eta\}]. \quad (7.11)$$

Step 2. Let us understand the law of $\{\ell_n\}_{1 \leq n \leq N}$ under $\tilde{\mathbb{P}}_N\{\cdot|\mathcal{D}\}$ defined by (7.7). In particular we prove that under the measure defined by (7.7) the recovery rates for the names that have defaulted are independent and that the default probabilities are independent as well.

For any $\{\psi\}_{1 \leq n \leq N} \subset \mathbb{R}$, we have that

$$\begin{aligned} &\mathbb{E}_N \left[\exp \left[\sqrt{-1} \sum_{n=1}^N \psi_n \ell_n + N\theta L_N \right] \middle| \mathcal{D} \right] \\ &= \prod_{n=1}^N \left\{ \Delta_n \int_{r \in [0,1]} \exp[(\sqrt{-1}\psi_n + \theta)(1-r)] \varrho_{N,n}(v_N(X), dr) + 1 - \Delta_n \right\}. \end{aligned}$$

Thus

$$\begin{aligned} & \tilde{\mathbb{E}}_N \left[\exp \left[\sqrt{-1} \sum_{n=1}^N \psi_n \ell_n \right] \middle| \mathcal{D} \right] \\ &= \prod_{n=1}^N \left\{ \Delta_n \int_{r \in [0,1]} \exp[\sqrt{-1} \psi_n (1-r)] \tilde{Q}_{N,n}(v_N(\mathbf{X}), dr) + 1 - \Delta_n \right\} \end{aligned}$$

where

$$\tilde{Q}_{N,n}(D, A) \stackrel{\text{def}}{=} \frac{\int_{r \in [0,1] \cap A} \exp[\theta(1-r)] Q_{N,n}(D, dr)}{\int_{r \in [0,1]} \exp[\theta(1-r)] Q_{N,n}(D, dr)} \quad A \in \mathcal{B}[0, 1], \quad D \in [0, 1]$$

for all $N \in \mathbb{N}$ and $n \in \{1, 2, \dots, N\}$. In other words, the recovery rates for the names which have defaulted are independent with laws given by the $\tilde{Q}_{N,n}(v_N(\mathbf{X}), \cdot)$'s. In particular,

$$\tilde{\mathbb{E}}_N [L_N | \mathcal{D}] = \frac{1}{N} \sum_{n=1}^N \Delta_n \frac{\int_{r \in [0,1]} (1-r) \exp[\theta(1-r)] Q_{N,n}(v_N(\mathbf{X}), dr)}{\int_{r \in [0,1]} \exp[\theta(1-r)] Q_{N,n}(v_N(\mathbf{X}), dr)} = \Lambda'_{v_N}(\theta).$$

Moreover,

$$\begin{aligned} & \tilde{\mathbb{E}}_N \left[|L_N - \Lambda_{v_N}(\theta)|^2 | \mathcal{D} \right] \\ &= \frac{1}{N^2} \sum_{n=1}^N \left\{ \int_{r \in [0,1]} (1-r)^2 \tilde{Q}_{N,n}(v_N(\mathbf{X}), dr) - \left(\int_{r \in [0,1]} (1-r) \tilde{Q}_{N,n}(v_N(\mathbf{X}), dr) \right)^2 \right\} \\ &\leq \frac{1}{N}. \end{aligned} \tag{7.12}$$

In a similar way, we next need to understand the statistics of the defaults under $\tilde{\mathbb{P}}_N$. For $\{\psi\}_{1 \leq n \leq N} \subset \mathbb{R}$,

$$\begin{aligned} & \mathbb{E}_N \left[\exp \left[\sqrt{-1} \sum_{n=1}^N \left\{ \psi_n \Delta_n + \phi(\gamma^{N,n}) \Delta_n \right\} \right] \right] \\ &= \prod_{n=1}^N \left\{ p^{N,n} \exp \left[\sqrt{-1} \psi_n + \phi(\gamma^{N,n}) \right] + 1 - p^{N,n} \right\}. \end{aligned}$$

Thus

$$\tilde{\mathbb{E}}_N \left[\exp \left[\sqrt{-1} \sum_{n=1}^N \psi_n \Delta_n \right] \right] = \prod_{n=1}^N \left\{ \tilde{p}^{N,n} \exp \left[\sqrt{-1} \psi_n \right] + 1 - \tilde{p}^{N,n} \right\} \tag{7.13}$$

where

$$\tilde{p}^{N,n} = \frac{p^{N,n} e^{\phi(\gamma^{N,n})}}{p^{N,n} e^{\phi(\gamma^{N,n})} + 1 - p^{N,n}} = \lambda'_{p^{N,n}}(\phi(\gamma^{N,n})) = \frac{dv^*}{dU}(\gamma^{N,n})$$

for all $N \in \mathbb{N}$ and $n \in \{1, 2, \dots, N\}$. In other words, under $\tilde{\mathbb{P}}_N$ the defaults are independent with probabilities given by the $\tilde{p}^{N,n}$'s.

Step 3. Let us go back to (7.11). We want to show that $\lim_{N \rightarrow \infty} \tilde{\mathbb{P}}_N(S_N) > 0$, which will in turn follow if $\overline{\lim}_{N \rightarrow \infty} \tilde{\mathbb{P}}_N(S_N^c) = 0$. To organize our thoughts, we write that

$$\begin{aligned} \tilde{\mathbb{P}}_N(S_N^c) &\leq \tilde{\mathbb{P}}_N\{v_N \notin \mathcal{O}_3\} + \tilde{\mathbb{E}}_N \left[\tilde{\mathbb{P}}_N\{|L_N - A'_{v_N}(\theta)| \geq \eta_2 | \mathcal{D}\} \chi_{\{v_N \in \mathcal{O}_3\}} \right] \\ &\leq \tilde{\mathbb{P}}_N\{v_N \notin \mathcal{O}_3\} + \frac{1}{\eta_2^2} \tilde{\mathbb{E}}_N \left[\tilde{\mathbb{E}}_N \left[|L_N - A'_{v_N}(\theta)|^2 | \mathcal{D} \right] \chi_{\{v_N \in \mathcal{O}_3\}} \right]. \end{aligned} \quad (7.14)$$

Let us show that

$$\lim_{N \rightarrow \infty} \tilde{\mathbb{E}}_N \left[\left| \int_{\gamma \in X} \Psi(\gamma) v_N(d\gamma) - \int_{\gamma \in X} \Psi(\gamma) v^*(d\gamma) \right| \right] = 0. \quad (7.15)$$

Fix $\Psi \in C(X)$. We write

$$\int_{\gamma \in X} \Psi(\gamma) v_N(d\gamma) - \int_{\gamma \in X} \Psi(\gamma) v^*(d\gamma) = \mathcal{E}_1^N + \mathcal{E}_2^N$$

where

$$\begin{aligned} \mathcal{E}_1^N &\stackrel{\text{def}}{=} \frac{1}{N} \sum_{n=1}^N \left\{ \Delta_n - \tilde{p}^{N,n} \right\} \Psi(\gamma^{N,n}) \\ \mathcal{E}_2^N &\stackrel{\text{def}}{=} \frac{1}{N} \sum_{n=1}^N \frac{dv^*}{dU}(\gamma^{N,n}) \Psi(\gamma^{N,n}) - \int_{\gamma \in X} \Psi(\gamma) v^*(d\gamma) \\ &= \int_{\gamma \in X} \frac{dv^*}{dU}(\gamma) \Psi(\gamma) U_N(d\gamma) - \int_{\gamma \in X} \frac{dv^*}{dU}(\gamma) \Psi(\gamma) U(d\gamma). \end{aligned}$$

From (7.13) we have that $\tilde{\mathbb{E}}_N[\mathcal{E}_1^N] = 0$; we also have by independence (Step 2) that

$$\tilde{\mathbb{E}}_N \left[\left| \mathcal{E}_1^N \right|^2 \right] \leq \frac{\sup_{\gamma \in X} |\Psi(\gamma)|^2}{N}.$$

The requirement that $\frac{dv^*}{dU} \in C(X)$ and Assumption 3.3 ensure that $\lim_{N \rightarrow \infty} \mathcal{E}_2^N = 0$.

Combining things together, we have (7.15). Since Ψ was an arbitrary element of $C(X)$ and X is Polish, (7.15) indeed implies (see [17])

$$\lim_{N \rightarrow \infty} \tilde{\mathbb{P}}_N\{v_N \notin \mathcal{O}_3\} = 0. \quad (7.16)$$

Using (7.12) and (7.16) we get by (7.14) that $\overline{\lim}_{N \rightarrow \infty} \tilde{\mathbb{P}}_N(S_N^c) = 0$. This and (7.11) give us the statement of the proposition. \square

We can now prove the full lower bound.

Proposition 7.3. *Let G be an open subset of $[0, 1] \times \mathcal{M}_1(X)$. Then*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \ln \mathbb{P}_N\{Z_N \in G\} \geq - \inf_{z \in G} I(z). \quad \square$$

Proof. Fix $z = (\ell, \nu) \in G$. If $I(z) < \infty$ and $\ell \in \text{ri dom } \Lambda_\nu^*$, then we get (7.1) from Proposition 7.2. If $I(z) = \infty$, we of course again get (7.1). Finally, assume that $\ell \in \text{dom } \Lambda_\nu^* \setminus$

$\text{ri dom } \Lambda_v^*$. We use the fact that $\text{dom } \Lambda_v^* \subset \overline{\text{ri dom } \Lambda_v^*}$ and convexity of $\ell \mapsto \Lambda_v^*(\ell)$. Fix a relaxation parameter $\eta > 0$. Then there is an $\ell' \in \text{ri dom } \Lambda_v^*$ such that $(\ell', v) \in G$ and $\Lambda_v^*(\ell') < \Lambda_v^*(\ell) + \eta$ (see [5, Appendix A]). Using Proposition 7.2, we get that

$$\lim_{N \rightarrow \infty} \frac{1}{N} \ln \mathbb{P}_N \{Z_N \in G\} \geq -I(\ell', v) \geq -I(z) - \eta.$$

Letting $\eta \searrow 0$, we again get (7.1). Letting z vary over G , we get the claim. \square

8. Large deviations upper bound

The heart of the upper bound is an exponential Chebyshev inequality. We will mimic, as much as possible, the proof of the upper bound of Cramér's theorem. The main result of this section is

Proposition 8.1. *Fix any closed subset F of $[0, 1] \times \mathcal{M}_1(\mathbf{X})$. Then*

$$\overline{\lim}_{N \rightarrow \infty} \frac{1}{N} \ln \mathbb{P}_N \{Z_N \in F\} \leq - \inf_{z \in F} I(z). \quad \square$$

Not surprisingly, we will first prove the bound for F compact; we will then show enough exponential tightness to get to the full claim.

Proposition 8.2. *Fix any compact subset F of $[0, 1] \times \mathcal{M}_1(\mathbf{X})$. Then*

$$\overline{\lim}_{N \rightarrow \infty} \frac{1}{N} \ln \mathbb{P}_N \{Z_N \in F\} \leq - \inf_{z \in F} I(z). \quad \square$$

Proof. To begin, fix $s < \inf_{z \in F} I(z)$. Fix also a relaxation parameter $\eta > 0$. For each $(\theta, \phi) \in \mathbb{R} \times C(\mathbf{X})$, define the set

$$\begin{aligned} \mathcal{O}_{(\theta, \phi)} &\stackrel{\text{def}}{=} \{(\ell, v) \in [0, 1] \times \mathcal{M}_1(\mathbf{X}) : \theta \ell \\ &\quad + \int_{\gamma \in \mathbf{X}} \phi(\gamma) v(d\gamma) - \Lambda_v(\theta) > s + \int_{\gamma=(p, \varrho) \in \mathbf{X}} \lambda_p(\phi(\gamma)) U(d\gamma)\} \end{aligned}$$

(these open sets were used in the proof of Proposition 6.1).

Fix now a $z \in F$. By definition of I and Lemma 4.7, we see that there is a $(\theta_z, \phi_z) \in \mathbb{R} \times C(\mathbf{X})$ such that $z \in \mathcal{O}_{(\theta_z, \phi_z)}$. Since $(\ell, v) \mapsto \theta \ell + \int_{\gamma \in \mathbf{X}} \phi_z(\gamma) v(d\gamma) - \Lambda_v(\theta)$ is continuous, there is an open neighborhood \mathcal{O}_z^* of z such that $\mathcal{O}_z^* \subset \mathcal{O}_{(\theta_z, \phi_z)}$ and such that

$$\theta_z \tilde{\ell} + \int_{\gamma \in \mathbf{X}} \phi_z(\gamma) \tilde{v}(d\gamma) - \Lambda_{\tilde{v}}(\theta_z) > s + \int_{\gamma=(p, \varrho) \in \mathbf{X}} \lambda_p(\phi_z(\gamma)) U(d\gamma)$$

for all $(\tilde{\ell}, \tilde{v}) \in \mathcal{O}_z^*$. Thus

$$F \subset \bigcup_{z \in F} \mathcal{O}_z^*,$$

the compactness of F implies that we can extract a finite subset \mathcal{Z} of F such that

$$F \subset \bigcup_{z \in \mathcal{Z}} \mathcal{O}_z^*$$

and thus

$$\mathbb{P}_N \{Z_N \in F\} \leq \sum_{z \in \mathcal{Z}} \mathbb{P}_N \{Z_N \in \mathcal{O}_z^*\}.$$

Fix now $z \in \mathcal{Z}$. We have that

$$\begin{aligned} \mathbb{P}_N \{Z_N \in \mathcal{O}_z^*\} &\leq \mathbb{P}_N \left\{ \theta_z L_N + \int_{\gamma \in X} \phi_z(\gamma) v_N(d\gamma) > s + \Lambda_{v_N}(\theta_z) \right. \\ &\quad \left. + \int_{\gamma=(p,q) \in X} \lambda_p(\phi_z(\gamma)) \mathbf{U}(d\gamma) \right\} \\ &\leq e^{-Ns} \mathbb{E}_N \left[\exp \left[N \left\{ \theta_z L_N - \Lambda_{v_N}(\theta_z) \right\} \right] \exp \left[N \left\{ \int_{\gamma \in X} \phi_z(\gamma) v_N(d\gamma) \right. \right. \right. \\ &\quad \left. \left. \left. - \int_{\gamma=(p,q) \in X} \lambda_p(\phi_z(\gamma)) \mathbf{U}(d\gamma) \right\} \right] \right] \\ &= e^{-Ns} \exp \left[N \left\{ \int_{\gamma=(p,q) \in X} \lambda_p(\phi_z(\gamma)) \mathbf{U}_N(d\gamma) \right. \right. \\ &\quad \left. \left. - \int_{\gamma=(p,q) \in X} \lambda_p(\phi_z(\gamma)) \mathbf{U}(d\gamma) \right\} \right]. \end{aligned}$$

We have used here the fact that

$$\mathbb{E}_N \left[\exp \left[N \left\{ \theta_z L_N - \Lambda_{v_N}(\theta_z) \right\} \right] \middle| \mathcal{D} \right] = 1$$

(compare with (7.6)) and that

$$\mathbb{E}_N \left[\exp \left[N \int_{\gamma \in X} \phi_z(\gamma) v_N(d\gamma) \right] \right] = \exp \left[N \int_{\gamma=(p,q) \in X} \lambda_p(\phi_z(\gamma)) \mathbf{U}_N(d\gamma) \right].$$

Letting $N \rightarrow \infty$, we get that

$$\overline{\lim}_{N \rightarrow \infty} \frac{1}{N} \ln \mathbb{P}_N \{Z_N \in \mathcal{O}_z^*\} \leq -s.$$

This gives the claim. \square

Let us next show that v_N is in a compact set.

Proposition 8.3 (Exponential Tightness). *For each $L > 0$ there is a compact subset \mathcal{K}_L of $\mathcal{M}_1(X)$ such that*

$$\overline{\lim}_{N \rightarrow \infty} \frac{1}{N} \ln \mathbb{P}_N \{v_n \notin \mathcal{K}_L\} \leq -L. \quad \square$$

Proof. First note that Assumption 3.3 implies that $\{\mathbf{U}_N\}_{N \in \mathbb{N}}$ is tight. Thus for each $j \in \mathbb{N}$, there is a compact subset K_j of X such that

$$\sup_{N \in \mathbb{N}} \mathbf{U}_N(X \setminus K_j) \leq \frac{1}{(L+j)^2}.$$

We will define

$$\mathcal{K}_L \stackrel{\text{def}}{=} \overline{\left\{ v \in \mathcal{M}_1(X) : v(X \setminus K_j) \leq \frac{1}{L+j} \text{ for all } j \in \mathbb{N} \right\}}.$$

Then \mathcal{K}_L is compact, and we have that

$$\begin{aligned}\mathbb{P}_N\{v_N \notin \mathcal{K}_L\} &\leq \sum_{j=1}^{\infty} \mathbb{P}_N\left\{v_N(\mathbf{X} \setminus K_j) \geq \frac{1}{L+j}\right\} \\ &\leq \sum_{j=1}^{\infty} \mathbb{P}_N\left\{N(L+j)^2 v_N(\mathbf{X} \setminus K_j) \geq N(L+j)\right\} \\ &\leq \sum_{j=1}^{\infty} \exp[-N(L+j)] \mathbb{E}_N\left[\exp\left[N(L+j)^2 v_N(\mathbf{X} \setminus K_j)\right]\right].\end{aligned}$$

We now compute that

$$\begin{aligned}\mathbb{E}_N\left[\exp\left[N(L+j)^2 v_N(\mathbf{X} \setminus K_j)\right]\right] &= \prod_{n=1}^N \mathbb{E}_N\left[\exp\left[N(L+j)^2 \sum_{n=1}^N \Delta_n \chi_{\mathbf{X} \setminus K_j}(\gamma^{N,n})\right]\right] \\ &= \exp\left[N \sum_{n=1}^N \lambda_{p^{N,n}}\left((L+j)^2 \chi_{\mathbf{X} \setminus K_j}(\gamma^{N,n})\right)\right] \\ &\leq \exp\left[N(L+j)^2 U_N(\mathbf{X} \setminus K_j)\right] \leq e^N.\end{aligned}$$

We have used here the calculation that for $\theta > 0$,

$$\lambda_p(\theta) \leq \ln(pe^\theta + (1-p)e^\theta) = \theta.$$

Combining things together, we get that

$$\begin{aligned}\mathbb{P}_N\{v_N \notin \mathcal{K}_L\} &\leq \sum_{j=1}^{\infty} e^{-N(L+j)} e^N = e^{-NL} \sum_{j=1}^{\infty} e^{-N(j-1)} \\ &\leq e^{-NL} \sum_{j=1}^{\infty} e^{-(j-1)} = \frac{e^{-NL}}{1-e^{-1}}. \quad \square\end{aligned}$$

We can now get the full upper bound.

Proof of Proposition 8.1. Fix $L > 0$. Then

$$\mathbb{P}_N\{Z_N \in F\} \leq \mathbb{P}_N\{Z_N \in F, v_N \in \mathcal{K}_L\} + \mathbb{P}_N\{v_N \notin \mathcal{K}_L\}.$$

We use [Proposition 8.2](#) on the first term together with the fact that $[0, 1] \times \mathcal{K}_L$ is compact. We use [Proposition 8.3](#) on the second term. Noting that (Lemma 1.2.15 in [5])

$$\overline{\lim}_{N \rightarrow \infty} \frac{1}{N} \ln(A_N + B_N) = \max\left\{\overline{\lim}_{N \rightarrow \infty} \frac{1}{N} \ln A_N, \overline{\lim}_{N \rightarrow \infty} \frac{1}{N} \ln B_N\right\}$$

we get

$$\overline{\lim}_{N \rightarrow \infty} \frac{1}{N} \ln \mathbb{P}_N\{Z_N \in F, v_N \in \mathcal{K}_L\} \leq - \inf_{\substack{z=(\ell, v) \in F \\ v \in \mathcal{K}_L}} I(z) \leq - \inf_{z \in F} I(z). \quad \square$$

9. Alternative expression for the rate function

In this section, we discuss the alternative expression for the rate function I' of [Theorem 4.6](#) given by [Theorem 4.8](#). In particular, this alternative representation shows that $I'(\ell)$ has a natural

interpretation as the favored way to rearrange recoveries and losses among the different types. In addition to providing intuitive insight, this alternative expression suggests numerical schemes for computing the rate function. We will rigorously verify that the alternative expression is correct, but will be heuristic in our discussion of the numerical schemes.

We defer the proof of [Theorem 4.8](#) to the end of this section and we first study the variational problem (4.12) using a Lagrange multiplier approach. Even though an explicit expression is usually not available, one can use numerical optimization techniques to calculate the quantities involved. In order to do that, we first recall that we can rewrite J' of (4.12) as a two-stage minimization problem, see expression (4.13).

This naturally suggests an analysis via a Lagrangian. Define

$$L(\Phi, \Psi, \lambda_1, \lambda_2) = \int_{\gamma=(p,\varrho)\in X} \left\{ \Phi(p) I_{\varrho}(\Psi(\gamma), D) + \hbar_p(\Phi(\gamma)) \right\} U(d\gamma) \\ - \lambda_1 \left\{ \int_{\gamma\in X} \Phi(\gamma) \Psi(\gamma) U(d\gamma) - \ell \right\} - \lambda_2 \left\{ \int_{\gamma\in X} \Phi(\gamma) U(d\gamma) - D \right\}.$$

Let us assume that Φ^* and Ψ^* are the minimizers. Let us also assume that $I_{\varrho}(\cdot, D)$ is differentiable for all $\gamma = (p, \varrho)$ in the support of U . We should then have that for every η_1 and η_2 in $\mathcal{B} = B(X; [0, 1])$,

$$\int_{\gamma=(p,\varrho)\in X} \eta_1(\gamma) \left\{ I_{\varrho}(\Psi^*(\gamma), D) + \hbar'_p(\Phi^*(\gamma)) - \lambda_1 \Psi^*(\gamma) - \lambda_2 \right\} U(d\gamma) = 0 \\ \int_{\gamma=(p,\varrho)\in X} \eta_2(\gamma) \Phi^*(\gamma) \left\{ I'_{\varrho}(\Psi^*(\gamma), D) - \lambda_1 \right\} U(d\gamma) = 0.$$

Ignoring any complications which would arise on the set where $\Phi^* = 0$, we should then have that

$$I_{\varrho}(\Psi^*(\gamma), D) + \hbar'_p(\Phi^*(\gamma)) = \lambda_1 \Psi^*(\gamma) + \lambda_2 \\ I'_{\varrho}(\Psi^*(\gamma), D) = \lambda_1$$

for all $\gamma = (p, \varrho) \in X$. This is a triangular system; the first equation depends on both λ_1 and λ_2 , but the second depends only on λ_1 . Recalling now (4.2) and the structure of Legendre–Fenchel transforms, we should have that

$$M'_{\varrho}(\lambda_1, D) = \Psi^*(\gamma) \\ \hbar'_p(\Phi^*(\gamma)) = \lambda_2 + \lambda_1 M'_{\varrho}(\lambda_1, D) - I_{\varrho}(\Psi^*(\gamma), D) = \lambda_2 + M_{\varrho}(\lambda_1, D)$$

for all $\gamma = (p, \varrho) \in X$. We can then invert this. This leads us to the following. Define

$$\Phi_{\lambda_1, \lambda_2, D}(p, \varrho) \stackrel{\text{def}}{=} \frac{p e^{\lambda_2 + M_{\varrho}(\lambda_1, D)}}{1 - p + p e^{\lambda_2 + M_{\varrho}(\lambda_1, D)}} \quad \lambda_1, \lambda_2 \in \mathbb{R}, (p, \varrho) \in X \\ \Psi_{\lambda_1, D}(\varrho) \stackrel{\text{def}}{=} M'_{\varrho}(\lambda_1, D) \quad \lambda_1 \in \mathbb{R}, \varrho \in C([0, 1]; \mathcal{P}[0, 1])$$

where $(\lambda_1, \lambda_2) = (\lambda_1(\ell, D, U), \lambda_2(\ell, D, U))$ is such that

$$\int_{\gamma\in X} \Phi_{\lambda_1, \lambda_2, D}(\gamma) U(d\gamma) = D \\ \int_{\gamma\in X} \Phi_{\lambda_1, \lambda_2, D}(\gamma) \Psi_{\lambda_1, \lambda_2, D}(\gamma) U(d\gamma) = \ell.$$

We conclude this section with the rigorous proof of the alternate representation.

Proof of Theorem 4.8. First, we prove that $J'(\ell) \geq I'(\ell)$. Consider any Φ and $\Psi \in \mathcal{B}$ such that

$$\int_{\gamma \in X} \Phi(\gamma) \Psi(\gamma) \mathbf{U}(d\gamma) = \ell. \quad (9.1)$$

For any $\theta \in \mathbb{R}$,

$$\begin{aligned} & \int_{\gamma=(p,q) \in X} \left\{ \Phi(\gamma) I_q \left(\Psi(\gamma), \int_{\gamma=(p,q) \in X} \Phi(\gamma) \mathbf{U}(d\gamma) \right) + \hbar_p(\Phi(\gamma)) \right\} \mathbf{U}(d\gamma) \\ &= \int_{\gamma=(p,q) \in X} \left\{ \Phi(\gamma) \sup_{\theta' \in \mathbb{R}} \left\{ \theta' \Psi(\gamma) - M_q \left(\theta', \int_{\gamma=(p,q) \in X} \Phi(\gamma) \mathbf{U}(d\gamma) \right) \right\} \right. \\ & \quad \left. + \hbar_p(\Phi(\gamma)) \right\} \mathbf{U}(d\gamma) \\ &\geq \int_{\gamma=(p,q) \in X} \left\{ \Phi(\gamma) \left\{ \theta \Psi(\gamma) - M_q \left(\theta, \int_{\gamma=(p,q) \in X} \Phi(\gamma) \mathbf{U}(d\gamma) \right) \right\} \right. \\ & \quad \left. + \hbar_p(\Phi(\gamma)) \right\} \mathbf{U}(d\gamma) \\ &= \theta \ell - \int_{\gamma=(p,q) \in X} \left\{ \Phi(\gamma) M_q \left(\theta, \int_{\gamma \in X} \Phi(\gamma) \mathbf{U}(d\gamma) \right) + \hbar_p(\Phi(\gamma)) \right\} \mathbf{U}(d\gamma). \end{aligned}$$

Define $\nu \in \mathcal{M}_1(X)$ as

$$\nu(A) \stackrel{\text{def}}{=} \int_{\gamma \in A} \Phi(\gamma) \mathbf{U}(d\gamma); \quad A \in \mathcal{B}(X)$$

then

$$\begin{aligned} & \int_{\gamma=(p,q) \in X} \left\{ \Phi(\gamma) I_q \left(\Psi(\gamma), \int_{\gamma \in X} \Phi(\gamma) \mathbf{U}(d\gamma) \right) + \hbar_p(\Phi(\gamma)) \right\} \mathbf{U}(d\gamma) \\ &\geq \theta \ell - \int_{\gamma=(p,q) \in X} M_q(\theta, \nu(X)) \nu(d\gamma) + \int_{\gamma=(p,q) \in X} \hbar_p \left(\frac{d\nu}{d\mathbf{U}}(\gamma) \right) \mathbf{U}(d\gamma). \end{aligned}$$

Varying θ , we get that

$$\begin{aligned} & \int_{\gamma=(p,q) \in X} \left\{ \Phi(\gamma) I_q \left(\Psi(\gamma), \int_{\gamma=(p,q) \in X} \Phi(\gamma) \mathbf{U}(d\gamma) \right) + \hbar_p(\Phi(\gamma)) \right\} \mathbf{U}(d\gamma) \\ &\geq \Lambda_v^*(\ell) + H(\nu) \geq I'(\ell) \end{aligned}$$

and then varying Φ and Ψ in \mathcal{B} (such that (9.1) holds), we get that $J'(\ell) \geq I'(\ell)$.

To show that $I'(\ell) \geq J'(\ell)$, fix $\nu \in \mathcal{M}_1(X)$ such that $\nu \ll \mathbf{U}$. We want to show that

$$\Lambda_v^*(\ell) + H(\nu) \geq J'(\ell). \quad (9.2)$$

For all $q \in C([0, 1]; \mathcal{P}[0, 1])$, define

$$\begin{aligned} \alpha_-(q, D) &\stackrel{\text{def}}{=} \inf\{1 - r \in [0, 1] : r \in \text{supp} q(D, \cdot)\} \\ \alpha_+(q, D) &\stackrel{\text{def}}{=} \sup\{1 - r \in [0, 1] : r \in \text{supp} q(D, \cdot)\}. \end{aligned}$$

Dominated convergence implies that

$$\lim_{\theta \rightarrow -\infty} \Lambda'_v(\theta) = \bar{\alpha}_- \stackrel{\text{def}}{=} \int_{\gamma=(p,q) \in X} \alpha_-(q, v(X)) v(d\gamma)$$

$$\lim_{\theta \rightarrow \infty} \Lambda'_v(\theta) = \bar{\alpha}_+ \stackrel{\text{def}}{=} \int_{\gamma=(p,q) \in X} \alpha_+(q, v(X)) v(d\gamma).$$

From (4.11) and the monotonicity of moment generating functions, we can see that Λ_v is nondecreasing; thus $(\bar{\alpha}_-, \bar{\alpha}_+) \in \Lambda'_v(\mathbb{R})$. This leads to three possible cases.

Case 1: Assume that $\ell \in (\bar{\alpha}_-, \bar{\alpha}_+)$, and let $\theta^* \in \mathbb{R}$ be such that $\Lambda'_v(\theta^*) = \ell$; i.e.,

$$\int_{\gamma=(p,q) \in X} M'_q(\theta^*, v(X)) v(d\gamma) = \ell. \quad (9.3)$$

Then

$$\begin{aligned} \Lambda_v^*(\ell) + H(v) &= \sup_{\theta \in \mathbb{R}} \left\{ \theta \ell - \int_{\gamma=(p,q) \in X} M_q(\theta, v(X)) v(d\gamma) \right\} + \int_{\gamma=(p,q) \in U} \hbar_p \left(\frac{dv}{dU}(\gamma) \right) U(d\gamma) \\ &\geq \theta^* \ell - \int_{\gamma=(p,q) \in X} M_q(\theta^*, v(X)) v(d\gamma) + \int_{\gamma=(p,q) \in U} \hbar_p \left(\frac{dv}{dU}(\gamma) \right) U(d\gamma) \\ &= \theta^* \Lambda'_v(\theta^*) - \int_{\gamma=(p,q) \in X} M_q(\theta^*, v(X)) v(d\gamma) + \int_{\gamma=(p,q) \in U} \hbar_p \left(\frac{dv}{dU}(\gamma) \right) U(d\gamma) \\ &= \int_{\gamma=(p,q) \in X} \left(\theta^* M'_q(\theta^*, v(X)) - M_q(\theta^*, v(X)) \right) v(d\gamma) \\ &\quad + \int_{\gamma=(p,q) \in U} \hbar_p \left(\frac{dv}{dU}(\gamma) \right) U(d\gamma) \\ &= \int_{\gamma=(p,q) \in X} I_q(M'_q(\theta^*, v(X)), v(X)) v(d\gamma) + \int_{\gamma=(p,q) \in U} \hbar_p \left(\frac{dv}{dU}(\gamma) \right) U(d\gamma). \end{aligned}$$

Define now $\Phi(\gamma) \stackrel{\text{def}}{=} \frac{dv}{dU}(\gamma)$ and $\Psi(\gamma) \stackrel{\text{def}}{=} M'_q(\theta^*, v(X))$. By (10.3) and (4.10) respectively we have that $\Phi, \Psi \in \mathcal{B}$. Then (9.3) is exactly that $\int_{\gamma \in X} \Phi(\gamma) \Psi(\gamma) U(d\gamma) = \ell$. Thus

$$\begin{aligned} \Lambda_v^*(\ell) + H(v) &\geq \int_{\gamma=(p,q) \in X} \left[\Phi(\gamma) I_q \left(\Psi(\gamma), \int_{\gamma \in X} \Phi(\gamma) U(d\gamma) \right) + \hbar_p(\Phi(\gamma)) \right] U(d\gamma) \\ &\geq J'(\ell). \end{aligned}$$

This is exactly (9.2).

Case 2: Assume next that $\ell \in [\bar{\alpha}_+, 1]$. For every $q \in C([0, 1]; \mathcal{P}[0, 1])$, define

$$\mathcal{E}_+^q(\theta) \stackrel{\text{def}}{=} M_q(\theta, v(X)) - \theta \alpha_+(q, v(X)) = \ln \int_{r \in [0, 1]} e^{-\theta(\alpha_+(q, v(X)) - (1-r))} q(v(X), dr)$$

for all $\theta \in \mathbb{R}$; thus

$$\begin{aligned} M_q(\theta, v(X)) &= \theta \alpha_+(q, v(X)) + \mathcal{E}_+^q(\theta) \quad q \in C([0, 1]; \mathcal{P}[0, 1]) \\ \Lambda_v(\theta) &= \theta \bar{\alpha}_+ + \int_{\gamma=(p,q) \in X} \mathcal{E}_+^q(\theta) v(d\gamma) \end{aligned}$$

for all $\theta \in \mathbb{R}$. For all $\varrho \in C([0, 1]; \mathcal{P}[0, 1])$ and $(1 - r) \in \text{supp} \varrho(v(\mathbf{X}), \cdot)$, the mapping $\theta \mapsto e^{-\theta(\alpha_+(\varrho, v(\mathbf{X})) - (1-r))}$ is decreasing and maps $[0, \infty)$ into $(0, 1]$. Monotone convergence implies that

$$\lim_{\theta \rightarrow \infty} \mathcal{E}_+^{\varrho}(\theta) = \ln \varrho\{v(\mathbf{X}), 1 - \alpha_+(\varrho, v(\mathbf{X}))\}.$$

If $\ell > \bar{\alpha}_+$, then we can use the fact that $\int_{\gamma=(p, \varrho) \in \mathbf{X}} \mathcal{E}_+^{\varrho}(\theta) v(d\gamma) \leq 0$ for all $\theta > 0$ to see that

$$\Lambda_v^*(\ell) \geq \overline{\lim}_{\theta \rightarrow \infty} \left\{ \theta(\ell - \bar{\alpha}_+) - \int_{\gamma=(p, \varrho) \in \mathbf{X}} \mathcal{E}_+^{\varrho}(\theta) v(d\gamma) \right\} \geq \overline{\lim}_{\theta \rightarrow \infty} \{\theta(\ell - \bar{\alpha}_+)\} = \infty \geq J'(\ell).$$

If $\ell = \bar{\alpha}_+$, then by the monotonicity of the \mathcal{E}_+^{ϱ} 's,

$$\begin{aligned} I_{\varrho}(\alpha_+(\varrho, v(\mathbf{X})), v(\mathbf{X})) &= \sup_{\theta \in \mathbb{R}} \{-\mathcal{E}_+^{\varrho}(\theta)\} = \overline{\lim}_{\theta \rightarrow \infty} \{-\mathcal{E}_+^{\varrho}(\theta)\} \\ &= \ln \left(\frac{1}{\varrho\{v(\mathbf{X}), 1 - \alpha_+(\varrho, v(\mathbf{X}))\}} \right) \end{aligned} \quad (9.4)$$

for all $\varrho \in C([0, 1]; \mathcal{P}[0, 1])$, and

$$\begin{aligned} \Lambda_v^*(\ell) &= \sup_{\theta \in \mathbb{R}} \left\{ - \int_{\gamma=(p, \varrho) \in \mathbf{X}} \mathcal{E}_+^{\varrho}(\theta) v(d\gamma) \right\} = \overline{\lim}_{\theta \rightarrow \infty} \left\{ - \int_{\gamma=(p, \varrho) \in \mathbf{X}} \mathcal{E}_+^{\varrho}(\theta) v(d\gamma) \right\} \\ &= \int_{\gamma=(p, \varrho) \in \mathbf{X}} \ln \left(\frac{1}{\varrho\{v(\mathbf{X}), 1 - \alpha_+(\varrho, v(\mathbf{X}))\}} \right) v(d\gamma) \\ &= \int_{\gamma=(p, \varrho) \in \mathbf{X}} I_{\varrho}(\alpha_+(\varrho, v(\mathbf{X}))) v(d\gamma). \end{aligned}$$

Defining $\Phi(\gamma) \stackrel{\text{def}}{=} \frac{dv}{d\mathbf{U}}(\gamma)$ and $\Psi(\gamma) = \alpha_+(\varrho, v(\mathbf{X}))$, we have that

$$\int_{\gamma \in \mathbf{X}} \Phi(\gamma) \Psi(\gamma) \mathbf{U}(d\gamma) = \int_{\gamma=(p, \varrho) \in \mathbf{X}} \alpha_+(\varrho, v(\mathbf{X})) v(d\gamma) = \ell.$$

Collecting things together, we see that if $\ell = \bar{\alpha}_+$, we again get (9.2).

Case 3: We finally assume that $\ell \in [0, \bar{\alpha}_-]$. The calculations are very similar to those of Case 2. For every $\varrho \in C([0, 1]; \mathcal{P}[0, 1])$, define

$$\mathcal{E}_-^{\varrho}(\theta) \stackrel{\text{def}}{=} M_{\varrho}(\theta, v(\mathbf{X})) - \theta \alpha_-(\varrho, v(\mathbf{X})) = \ln \int_{r \in [0, 1]} e^{\theta((1-r) - \alpha_-(\varrho, v(\mathbf{X})))} \varrho(v(\mathbf{X}), dr)$$

for all $\theta \in \mathbb{R}$, so that

$$\begin{aligned} M_{\varrho}(\theta, v(\mathbf{X})) &= \theta \alpha_-(\varrho, v(\mathbf{X})) + \mathcal{E}_-^{\varrho}(\theta) \varrho \in C([0, 1]; \mathcal{P}[0, 1]) \\ \Lambda_v(\theta) &= \theta \bar{\alpha}_- + \int_{\gamma=(p, \varrho) \in \mathbf{X}} \mathcal{E}_-^{\varrho}(\theta) v(d\gamma) \end{aligned}$$

for all $\theta \in \mathbb{R}$. For all $\varrho \in C([0, 1]; \mathcal{P}[0, 1])$ and $(1 - r) \in \text{supp} \varrho(v(\mathbf{X}), \cdot)$, the mapping $\theta \mapsto e^{\theta((1-r) - \alpha_-(\varrho, v(\mathbf{X})))}$ is increasing and maps $(-\infty, 0]$ into $(0, 1]$. Monotone convergence implies that

$$\lim_{\theta \rightarrow -\infty} \mathcal{E}_-^{\varrho}(\theta) = \ln \varrho\{v(\mathbf{X}), 1 - \alpha_-(\varrho, v(\mathbf{X}))\}.$$

If $\ell < \bar{\alpha}_-$, then we can use the fact that $\int_{\gamma=(p,\varrho)\in X} \mathcal{E}_-^\varrho(\theta) v(d\gamma) \geq 0$ for all $\theta < 0$ to see that

$$\begin{aligned} A_v^*(\ell) &\geq \overline{\lim}_{\theta \rightarrow -\infty} \left\{ \theta(\ell - \bar{\alpha}_-) - \int_{\gamma=(p,\varrho)\in X} \mathcal{E}_-^\varrho(\theta) v(d\gamma) \right\} \\ &\geq \overline{\lim}_{\theta \rightarrow -\infty} \{\theta(\ell - \bar{\alpha}_-)\} = \infty \geq J'(\ell). \end{aligned}$$

If $\ell = \bar{\alpha}_-$, then by the monotonicity of the \mathcal{E}_-^ϱ 's,

$$I_\varrho(\alpha_-(\varrho, v(X)), v(X)) = \sup_{\theta \in \mathbb{R}} \{-\mathcal{E}_-^\varrho(\theta)\} = \lim_{\theta \rightarrow -\infty} \{-\mathcal{E}_-^\varrho(\theta)\} = \ln \left(\frac{1}{\varrho\{1 - \alpha_-(\varrho, v(X))\}} \right)$$

for all $\varrho \in C([0, 1]; \mathcal{P}[0, 1])$, and

$$\begin{aligned} A_v^*(\ell) &= \sup_{\theta \in \mathbb{R}} \left\{ - \int_{\gamma=(p,\varrho)\in X} \mathcal{E}_-^\varrho(\theta) v(d\gamma) \right\} = \lim_{\theta \rightarrow -\infty} \left\{ - \int_{\gamma=(p,\varrho)\in X} \mathcal{E}_-^\varrho(\theta) v(d\gamma) \right\} \\ &= \int_{\gamma=(p,\varrho)\in X} \ln \left(\frac{1}{\varrho\{v(X), 1 - \alpha_-(\varrho, v(X))\}} \right) v(d\gamma) \\ &= \int_{\gamma=(p,\varrho)\in X} I_\varrho(\alpha_-(\varrho, v(X))) v(d\gamma). \end{aligned}$$

Defining $\Phi(\gamma) \stackrel{\text{def}}{=} \frac{dv}{dU}(\gamma)$ and $\Psi(\gamma) = \alpha_-(\varrho, v(X))$, we have that

$$\int_{\gamma \in X} \Phi(\gamma) \Psi(\gamma) U(d\gamma) = \int_{\gamma=(p,\varrho)\in X} \alpha_-(\varrho, v(X)) v(d\gamma) = \ell$$

again implying (9.2).

Collecting things together, we have (4.12). We get (4.13) by defining $D \stackrel{\text{def}}{=} \int_{\gamma \in X} \Phi(\gamma) U(d\gamma)$. Note that since Φ and Ψ both take values in $[0, 1]$,

$$\int_{\gamma \in X} \Psi(\gamma) \Phi(\gamma) U(d\gamma) \leq \int_{\gamma \in X} \Phi(\gamma) U(d\gamma).$$

This allows us to restrict the minimization in D to the interval $[\ell, 1]$. \square

10. Detailed structure of H

In this section we prove Lemmas 7.1 and 4.7. The discussion in this section is somewhat technical.

Recall the quantities defined in Definition 4.2. Note that

$$\hbar_0(x) = \begin{cases} 0 & \text{if } x = 0 \\ \infty & \text{else} \end{cases} \quad \text{and} \quad \hbar_1(x) = \begin{cases} 0 & \text{if } x = 1 \\ \infty & \text{else} \end{cases} \quad (10.1)$$

Thus if $H(v) < \infty$, we can restrict the region of integration to get that

$$\int_{\gamma=(p,\varrho)\in X} \chi_{\{0,1\}}(p) \hbar_p \left(\frac{dv}{dU}(\gamma) \right) U(d\gamma) < \infty$$

so in fact

$$\begin{aligned} \mathbb{U} \left\{ \gamma = (p, \varrho) \in \mathbf{X} : p = 0 \text{ and } \frac{dv}{d\mathbb{U}}(\gamma) \neq 0 \right\} &= 0 \\ \mathbb{U} \left\{ \gamma = (p, \varrho) \in \mathbf{X} : p = 1 \text{ and } \frac{dv}{d\mathbb{U}}(\gamma) \neq 1 \right\} &= 0. \end{aligned} \quad (10.2)$$

Fix $v \in \mathcal{M}_1(\mathbf{X})$ such that $H(v) < \infty$. The main technical challenges in both proofs is to stay away from the singularities in \hbar_p and \hbar'_p . Note that

$$\hbar'_p(x) = \ln \left(\frac{x}{1-x} \frac{1-p}{p} \right), \quad x, p \in (0, 1)$$

and keeping (10.1) in mind, we thus need to be careful near $p \in \{0, 1\}$, and for $(x, p) \in \{0, 1\} \times (0, 1)$.

To start, let us note some implications of the assumption that $H(v) < \infty$. Clearly $v \ll \mathbb{U}$. Second,

$$\mathbb{U} \left\{ \gamma \in \mathbf{X} : \frac{dv}{d\mathbb{U}}(\gamma) > 1 \right\} = 0. \quad (10.3)$$

Let us now do the following. Fix $N \in \mathbb{N}$. Define

$$\xi_N(\gamma) \stackrel{\text{def}}{=} \begin{cases} p & \text{if } p \notin \left[\frac{1}{N}, 1 - \frac{1}{N} \right] \\ \frac{dv}{d\mathbb{U}}(\gamma) & \text{if } p \in \left[\frac{1}{N}, 1 - \frac{1}{N} \right] \text{ and } \frac{dv}{d\mathbb{U}}(\gamma) \in \left(\frac{1}{N}, 1 - \frac{1}{N} \right) \\ \frac{1}{N} & \text{if } p \in \left[\frac{1}{N}, 1 - \frac{1}{N} \right] \text{ and } \frac{dv}{d\mathbb{U}}(\gamma) \leq \frac{1}{N} \\ 1 - \frac{1}{N} & \text{if } p \in \left[\frac{1}{N}, 1 - \frac{1}{N} \right] \text{ and } \frac{dv}{d\mathbb{U}}(\gamma) \geq 1 - \frac{1}{N}. \end{cases} \quad (10.4)$$

Clearly $0 \leq \xi_N \leq 1$, so we can define $\nu_N \in \mathcal{M}_1(\mathbf{X})$ as

$$\nu_N(A) \stackrel{\text{def}}{=} \int_{\gamma \in A} \xi_N(\gamma) \mathbb{U}(d\gamma) \quad A \in \mathcal{B}(\mathbf{X}).$$

In light of (10.3) and (10.2), $\lim_{N \rightarrow \infty} \xi_N = \frac{dv}{d\mathbb{U}}$ \mathbb{U} -a.s., so it follows that $\lim_{N \rightarrow \infty} \nu_N = v$. We next compute that

$$\hbar_p(\xi_N(\gamma)) = \begin{cases} 0 & \text{if } p \notin \left[\frac{1}{N}, 1 - \frac{1}{N} \right] \\ \hbar_p \left(\frac{dv}{d\mathbb{U}}(\gamma) \right) & \text{if } \frac{1}{N} \leq p \leq 1 - \frac{1}{N} \text{ and } \frac{dv}{d\mathbb{U}}(\gamma) \in \left(\frac{1}{N}, 1 - \frac{1}{N} \right) \\ \hbar_p \left(\frac{1}{N} \right) & \text{if } \frac{1}{N} \leq p \leq 1 - \frac{1}{N} \text{ and } \frac{dv}{d\mathbb{U}}(\gamma) \leq \frac{1}{N} \\ \hbar_p \left(1 - \frac{1}{N} \right) & \text{if } \frac{1}{N} \leq p \leq 1 - \frac{1}{N} \text{ and } \frac{dv}{d\mathbb{U}}(\gamma) \geq 1 - \frac{1}{N}. \end{cases}$$

Using again (10.3) and (10.2), we have that $\lim_{N \rightarrow \infty} \bar{h}_p(\xi_N(\gamma)) = \bar{h}_p\left(\frac{dv}{dU}(\gamma)\right)$ for U -almost-all $\gamma = (p, \varrho) \in X$. If $p \in \left[\frac{1}{N}, 1 - \frac{1}{N}\right]$, then \bar{h}_p is increasing on $[p, 1] \supset \left[1 - \frac{1}{N}, 1\right]$ and decreasing on $[0, p] \supset \left[0, \frac{1}{N}\right]$. Thus $\bar{h}_p(\xi_N(\gamma)) \leq \bar{h}_p\left(\frac{dv}{dU}(\gamma)\right)$ for U -almost-all $\gamma = (p, \varrho) \in X$. Dominated convergence thus implies that $\lim_{N \rightarrow \infty} H(v_N) = H(v)$.

Proof of Lemma 7.1. Fix $N \in \mathbb{N}$; we want to approximate ξ_N by “nice” elements of $C(X)$. Note that

$$\xi_N(\gamma) = p\chi_{[0,1] \setminus [N^{-1}, 1-N^{-1}]}(p) + \chi_{[N^{-1}, 1-N^{-1}]}(p)\xi_N(\gamma).$$

Since U is regular (recall that X is Polish), we can approximate $\gamma = (p, \varrho) \mapsto \chi_{[N^{-1}, 1-N^{-1}]}(p)\xi_N(\gamma)$ by elements of $C(X)$. From (10.4), we have that $N^{-1} \leq \xi_N(\gamma) \leq 1 - N^{-1}$ if $\gamma = (p, \varrho) \in X$ is such that $N^{-1} \leq p \leq 1 - N^{-1}$, so we can truncate these approximations at N^{-1} and $1 - N^{-1}$ without any loss. Namely, there is a sequence $(\xi^{1,\varepsilon})_{\varepsilon>0}$ in $C(X)$ such that

$$N^{-1} \leq \xi^{1,\varepsilon} \leq 1 - N^{-1}$$

$$\lim_{\varepsilon \searrow 0} \int_{\gamma=(p,\varrho) \in X} \left| \xi^{1,\varepsilon}(\gamma) - \chi_{[N^{-1}, 1-N^{-1}]}(p)\xi_N(\gamma) \right| U(d\gamma) = 0. \quad (10.5)$$

For each $\varepsilon > 0$, let $\varphi_\varepsilon \in C([0, 1]; [0, 1])$ be such that $\varphi_\varepsilon(u) = 1$ if $u \in [N^{-1}, 1 - N^{-1}]$ and $\varphi_\varepsilon(u) = 0$ if $u \in [0, 1] \setminus [N^{-1} - \varepsilon, 1 - N^{-1} + \varepsilon]$. For each $\varepsilon > 0$, define

$$\xi^{2,\varepsilon}(\gamma) \stackrel{\text{def}}{=} p\{1 - \varphi_\varepsilon(p)\} + \xi^{1,\varepsilon}(\gamma)\varphi_\varepsilon(p)$$

for all $\gamma = (p, \varrho) \in X$. Then $\xi^{2,\varepsilon} \in C(X)$ for all $\varepsilon > 0$. We also have that

$$\begin{aligned} & \int_{\gamma=(p,\varrho) \in X} \left| \xi^{2,\varepsilon}(\gamma) - \xi_N(\gamma) \right| U(d\gamma) \\ & \leq U\left\{ \gamma = (p, \varrho) \in X : p \in [N^{-1} + \varepsilon, 1 - N^{-1} + \varepsilon] \setminus [N^{-1}, 1 - N^{-1}] \right\} \\ & \quad + \int_{\gamma=(p,\varrho) \in X} \chi_{[N^{-1}, 1-N^{-1}]}(p) \left| \xi^{1,\varepsilon}(\gamma) - \xi_N(\gamma) \right| U(d\gamma). \end{aligned}$$

Dominated convergence and (10.5) then ensure that

$$\lim_{\varepsilon \rightarrow 0} \int_{\gamma=(p,\varrho) \in X} \left| \xi^{2,\varepsilon}(\gamma) - \xi_N(\gamma) \right| U(d\gamma) = 0. \quad (10.6)$$

Clearly $0 \leq \xi^{2,\varepsilon} \leq 1$, so we can define $\nu_{N,\varepsilon} \in \mathcal{M}_1(X)$ as

$$\nu_{N,\varepsilon}(A) \stackrel{\text{def}}{=} \int_{\gamma \in A} \xi^{2,\varepsilon}(\gamma) U(d\gamma). \quad A \in \mathcal{B}(X).$$

Thanks to (10.6), we have that $\lim_{\varepsilon \rightarrow 0} \nu_{N,\varepsilon} = \nu_N$. Note next that for $\gamma = (p, \varrho) \in X$ such that $p \in [0, 1] \setminus [N^{-1} - \varepsilon, 1 - N^{-1} + \varepsilon]$,

$$\bar{h}_p(\xi^{2,\varepsilon}(\gamma)) - \bar{h}_p(\xi_N(\gamma)) = \bar{h}_p(p) - \bar{h}_p(p) = 0.$$

If $\gamma = (p, \varrho) \in X$ is such that $p \in [N^{-1} - \varepsilon, 1 - N^{-1} + \varepsilon]$, then

$$N^{-1} - \varepsilon \leq \xi^{2,\varepsilon}(\gamma) \leq 1 - N^{-1} + \varepsilon, \quad (10.7)$$

so if $\varepsilon < 1/(2N)$,

$$\left| \hbar_p(\xi^{2,\varepsilon}(\gamma)) - \hbar_p(\xi_N(\gamma)) \right| \leq \kappa \left| \xi^{2,\varepsilon}(\gamma) - \xi_N(\gamma) \right|$$

where

$$\kappa \stackrel{\text{def}}{=} \sup \left\{ |\hbar'_p(x)| : \frac{1}{2N} \leq x \leq 1 - \frac{1}{2N} \text{ and } \frac{1}{2N} \leq p \leq 1 - \frac{1}{2N} \right\}.$$

Thus if $\varepsilon < \frac{1}{2N}$,

$$\left| \hbar_p(\xi^{2,\varepsilon}(\gamma)) - \hbar_p(\xi_N(\gamma)) \right| \leq \kappa \left| \xi^{2,\varepsilon}(\gamma) - \xi_N(\gamma) \right|$$

for all $\gamma = (p, \varrho) \in \mathbf{X}$. Thanks to (10.6), we thus have that $\lim_{\varepsilon \rightarrow 0} H(v_{N,\varepsilon}) = H(v)$.

We finally note that $\gamma = (p, \varrho) \mapsto \hbar'_p(\xi^{2,\varepsilon}(\gamma))$ is continuous on $\{\gamma = (p, \varrho) \in \mathbf{X} : p \in (N^{-1} - \varepsilon, 1 - N^{-1} + \varepsilon)\}$ ((10.7) ensures that $\tilde{\xi}_2$ takes values in $(0, 1)$ in this case). On $\{\gamma = (p, \varrho) \in \mathbf{X} : p \in (0, 1) \setminus (N^{-1} - \varepsilon, 1 - N^{-1} + \varepsilon)\}$, we have that $\hbar'_p(\xi^{2,\varepsilon}(\gamma)) = \hbar'_p(p) = 0$. This finishes the proof. \square

Proof of Lemma 4.7. Assume first that ν is not absolutely continuous with respect to \mathbf{U} . We will show that then the right-hand side of (4.9) is infinite. Then there is an $A \in \mathcal{B}(\mathbf{X})$ such that $\nu(A) > 0$ and $\mathbf{U}(A) = 0$. Since \mathbf{X} is Polish, ν is regular; i.e.,

$$\nu(A) = \sup \{ \nu(F) : F \subset A, F \text{ closed} \}.$$

Thus there is a closed subset F of A such that $\nu(F) > 0$. Fix also now $c > 0$. For each $n \in \mathbb{N}$, define

$$\phi_n(\gamma) \stackrel{\text{def}}{=} c \exp[-n \text{dist}(\gamma, F)] \quad \gamma \in \mathbf{X}$$

where $\text{dist}(\gamma, F)$ is the distance (in \mathbf{X}) from x to F . Then $0 \leq \phi_n \leq c$ for all $n \in \mathbb{N}$, and $\phi_n \searrow c\chi_F$. Recall the definition of $\lambda_p(\theta)$ from (4.8). Since $\theta \mapsto \lambda_p(\theta)$ is nondecreasing and continuous for each $p \in [0, 1]$, we also have that $\lambda_p(\phi_n(\gamma)) \searrow \lambda_p(c\chi_F(\gamma))$ for all $\gamma = (p, \varrho) \in \mathbf{X}$. Thus

$$\begin{aligned} & \sup_{\phi \in C(\mathbf{X})} \left\{ \int_{\gamma \in \mathbf{X}} \phi(\gamma) \nu(d\gamma) - \int_{\gamma=(p,\varrho) \in \mathbf{X}} \lambda_p(\phi(\gamma)) \mathbf{U}(d\gamma) \right\} \\ & \geq \overline{\lim}_{n \rightarrow \infty} \left\{ \int_{\gamma \in \mathbf{X}} \phi_n(\gamma) \nu(d\gamma) - \int_{\gamma=(p,\varrho) \in \mathbf{X}} \lambda_p(\phi_n(\gamma)) \mathbf{U}(d\gamma) \right\} = c\nu(F). \end{aligned}$$

Let $c \nearrow \infty$ to see that the right-hand side of (4.9) is infinite.

Assume next that $\nu \ll \mathbf{U}$. We use the fact that \hbar_p and λ_p are convex duals of each other. For any $\phi \in C(\mathbf{X})$,

$$\begin{aligned} & \int_{\gamma \in \mathbf{X}} \phi(\gamma) \nu(d\gamma) - \int_{\gamma=(p,\varrho) \in \mathbf{X}} \lambda_p(\phi(\gamma)) \mathbf{U}(d\gamma) \\ & = \int_{\gamma \in \mathbf{X}} \inf_{x \in \mathbb{R}} \left\{ \phi(\gamma) \left(\frac{d\nu}{d\mathbf{U}}(\gamma) - x \right) + \hbar_p(x) \right\} \mathbf{U}(d\gamma) \\ & \leq \int_{\gamma=(p,\varrho) \in \mathbf{X}} \hbar_p \left(\frac{d\nu}{d\mathbf{U}}(\gamma) \right) \mathbf{U}(d\gamma). \end{aligned}$$

To show the reverse inequality, let us write that

$$H(v) = \int_{\gamma=(p,q) \in X} \sup_{\theta \in \mathbb{R}} \left\{ \theta \frac{dv}{dU}(\gamma) - \lambda_p(\theta) \right\} U(d\gamma) = \int_{\gamma=(p,q) \in X} \lim_{N \rightarrow \infty} F_N(\gamma) U(d\gamma)$$

where

$$F_N(\gamma) = \sup_{|\theta| \leq N} \left\{ \theta \frac{dv}{dU}(\gamma) - \lambda_p(\theta) \right\} \quad (10.8)$$

for all $N \in \mathbb{N}$ and $\gamma = (p, q) \in X$. We can explicitly solve this minimization problem; for $N \in \mathbb{N}$ and $\gamma = (p, q) \in X$, define

$$\phi_N(\gamma) = \begin{cases} h'_p \left(\frac{dv}{dU}(\gamma) \right) & \text{if } p \in (0, 1), \frac{dv}{dU}(\gamma) \in (0, 1), \text{ and } -N \leq h'_p \left(\frac{dv}{dU}(\gamma) \right) \leq N \\ 0 & \text{if } p \in \{0, 1\} \text{ and } \frac{dv}{dU}(\gamma) = p \\ N \operatorname{sgn} \left(\frac{dv}{dU}(\gamma) - p \right) & \text{else} \end{cases}$$

(where sgn is the standard signum function). Then

$$F_N(\gamma) = \phi_N(\gamma) \frac{dv}{dU}(\gamma) - \lambda_p(\phi_N(\gamma))$$

for all $\gamma = (p, q) \in X$ and $N \in \mathbb{N}$. Clearly F_N and ϕ_N are measurable, and $\phi_N \in B(X)$. From (10.8), we also see that F_N is nondecreasing in N . Thus by (10.8) and monotone convergence

$$\begin{aligned} H(v) &= \lim_{N \rightarrow \infty} \int_{\gamma \in X} F_N(\gamma) U(d\gamma) = \lim_{N \rightarrow \infty} \int_{\gamma=(p,q) \in X} \left\{ \phi_N(\gamma) \frac{dv}{dU}(\gamma) - \lambda_p(\phi_N(\gamma)) \right\} U(d\gamma) \\ &\leq \sup_{\phi \in B(X)} \left\{ \int_{\gamma=(p,q) \in X} \phi(\gamma) v(d\gamma) - \int_{\gamma=(p,q) \in X} \lambda_p(\phi(\gamma)) U(d\gamma) \right\}. \end{aligned}$$

Since X is Polish, v and U are regular; and thus we can approximate elements of $B(X)$ by elements of $C(X)$, completing the proof. \square

Acknowledgment

The research of R.S. was supported in part by the National Science Foundation EFRI 1024837.

References

- [1] Edward I. Altman, Brooks Brady, Andrea Resti, Andrea Sironi, The link between default and recovery rates: theory, empirical evidence, and implications, *Journal of Business* 78 (6) (2005) 2203–2227.
- [2] Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, Heath David, Coherent measures of risk, *Mathematical Finance* 9 (3) (1999) 203–228.
- [3] Duffie Darell, Kenneth J. Singleton, *Credit Risk: Pricing, Measurement, and Management*, Princeton University Press, 2003.
- [4] Amir Dembo, Jean-Dominique Deuschel, Darrell Duffie, Large portfolio losses, *Finance and Stochastics* 8 (1) (2004) 3–16.
- [5] Amir Dembo, Ofer Zeitouni, *Large Deviations Techniques and Applications*, Springer-Verlag, New York, 1998.
- [6] Rudiger Frey, Alexander J. McNeil, VaR and expected shortfall in portfolios of dependent credit risks: conceptual and practical insights, *Journal of Banking & Finance* 26 (2002) 1317–1334.

- [7] Rudiger Frey, Alexander J. McNeil, Dependent defaults in models of portfolio credit risk, *Journal of Risk* 6 (1) (2003) 311–329.
- [8] Paul Glasserman, Jinyi Li, Importance sampling for portfolio credit risk, *Management Science* 51 (2005) 1643–1656.
- [9] Michael B. Gordy, Saddle point approximation of creditrisk, *Journal of Business Finance* 26 (2002) 1335–1353.
- [10] Michael B. Gordy, A risk-factor model foundation for ratings-based bank capital rules, *Journal of Financial Intermediation* 12 (3) (2003) 199–232.
- [11] Vincent Leijdekker, Michel Mandjes, Peter Spreij, Sample-path large deviations in credit risk, 2009. [arXiv:0909.5610v1](https://arxiv.org/abs/0909.5610v1) [math.PR].
- [12] Alexander J. McNeil, Rudiger Frey, Paul Embrechts, *Quantitative Risk Management: Concepts, Techniques and Tools*, Princeton University Press, 2005.
- [13] Glasserman Paul, Wanmo Kang, Perwez Shahabuddin, Large deviations in multifactor portfolio credit risk, *Mathematical Finance* 17 (3) (2007) 345–379.
- [14] Huy  n Pham, Some applications and methods of large deviations in finance and insurance, in: *Paris-Princeton Lectures on Mathematical Finance 2004*, in: *Lecture Notes in Math.*, vol. 1919, Springer, Berlin, 2007, pp. 191–244.
- [15] Das Sanjiv, Paul Hanouna, Implied recovery, *Journal of Economic Dynamics and Control* 33 (11) (2009) 1837–1857.
- [16] Richard B. Sowers, Exact pricing asymptotics of investment-grade tranches of synthetic cdo’s part i: a large homogeneous pool, *International Journal of Theoretical and Applied Finance* 13 (3) (2010) 367–401.
- [17] Daniel W. Stroock, *Probability Theory, An Analytic View*, Cambridge University Press, Cambridge, 1993.